

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 15:46:52 ; Search time 137 Seconds
(without alignments)

578.133 Million cell updates/sec

Title: US-10-620-099-2

Perfect score: 1270

Sequence: 1 MSKGEELTAVVILVELDGD.....VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1270	100.0	238	10	US-09-575-847-2
2	1270	100.0	238	14	US-10-071-976-2
3	1270	100.0	238	15	US-10-620-099-2
4	1270	100.0	238	17	US-10-924-232-2
5	1246	98.1	238	9	US-09-884-681-2
6	1246	98.1	238	10	US-09-967-301-2
7	1246	98.1	238	13	US-10-024-686-2
8	1246	98.1	238	13	US-10-057-505-2
9	1246	98.1	238	14	US-10-293-580-2
10	1246	98.1	238	14	US-10-293-580-74
11	1246	98.1	238	15	US-10-457-982-2
12	1246	98.1	238	16	US-10-724-178-2
13	1246	98.1	238	16	US-10-757-624-2

14	1242	97.8	238	9	US-09-920-922-4	Sequence 4, Appli
15	1242	97.8	238	9	US-09-852-000-1	Sequence 1, Appli
16	1242	97.8	238	10	US-09-900-345A-125	Sequence 135, App
17	1242	97.8	238	10	US-09-866-538-2	Sequence 2, Appli
18	1242	97.8	238	10	US-09-794-308-2	Sequence 2, Appli
19	1242	97.8	238	10	US-09-865-291-2	Sequence 2, Appli
20	1242	97.8	238	14	US-10-121-258-10	Sequence 10, Appli
21	1242	97.8	238	14	US-10-221-461-6	Sequence 6, Appli
22	1242	97.8	238	14	US-10-305-765-10	Sequence 10, Appli
23	1242	97.8	238	14	US-10-305-765-159	Sequence 159, App
24	1242	97.8	238	14	US-10-305-633-10	Sequence 10, Appli
25	1242	97.8	238	14	US-10-305-633-159	Sequence 159, App
26	1242	97.8	238	15	US-10-132-067-8	Sequence 8, Appli
27	1242	97.8	238	15	US-10-370-570-1	Sequence 1, Appli
28	1242	97.8	238	15	US-10-370-570-53	Sequence 53, Appli
29	1242	97.8	238	15	US-10-423-688A-42	Sequence 42, Appli
30	1242	97.8	238	15	US-10-568-168-4	Sequence 4, Appli
31	1242	97.8	238	16	US-10-333-680-3	Sequence 3, Appli
32	1242	97.8	238	17	US-10-857-622-2	Sequence 2, Appli
33	1241	97.7	238	16	US-10-676-428-35	Sequence 35, Appli
34	1239	97.6	238	15	US-10-668-168-2	Sequence 2, Appli
35	1239	97.6	243	10	US-09-900-345A-60	Sequence 60, Appli
36	1239	97.6	243	10	US-09-900-345A-62	Sequence 62, Appli
37	1239	97.6	243	10	US-09-900-345A-64	Sequence 64, Appli
38	1239	97.6	243	10	US-09-900-345A-66	Sequence 66, Appli
39	1239	97.6	243	10	US-09-900-345A-68	Sequence 68, Appli
40	1239	97.6	243	10	US-09-900-345A-70	Sequence 70, Appli
41	1239	97.6	243	14	US-10-305-765-94	Sequence 94, Appli
42	1239	97.6	243	14	US-10-305-765-96	Sequence 96, Appli
43	1239	97.6	243	14	US-10-305-765-98	Sequence 98, Appli
44	1239	97.6	243	14	US-10-305-765-100	Sequence 100, App
45	1239	97.6	243	14	US-10-305-765-102	Sequence 102, App

ALIGNMENTS

RESULT 1

US-09-575-847-2
; Sequence 2, Application US/09575847
; Publication No. US20030013149A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: WACHTER, Rebekka
; APPLICANT: REMINGTON, James
; TITLE OF INVENTION: LONG WAVELENGTH ENGINEERED FLUORESCENT PROTEINS
; FILE REFERENCE: REG1250-5
; CURRENT APPLICATION NUMBER: US/09/575,847
; CURRENT FILING DATE: 2000-05-19
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: US 08/974,737
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: US 08/911,825
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 08/706,408
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea
US-09-575-847-2

Query Match	100.0%	Score 1270;	DB 10;	Length 238;
Best Local Similarity	100.0%	Pred. No. 1.8e-119;	Indels 0;	Gaps 0;
Matches 238;	Conservative 0;	Mismatches 0;		
Oy	1	MSKGEELTAVVILVELDGDVNGHKFSVSGEGSDVTYGLTKLFICTTCKLPVPWPTL	60	
Db	1	MSKGEELTAVVILVELDGDVNGHKFSVSGEGSDVTYGLTKLFICTTCKLPVPWPTL	60	
Oy	61	VTTFSYGVQCFSRYPDHMKRHDFPKSAMPEGYVOORTIFFKDDGNKYKTRAEVKEGDTLV	120	

Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YQOQNTPLDGPVLLPDNHYLSQTSALSQKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YQOQNTPLDGPVLLPDNHYLSQTSALSQKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 2

US-10-071-976-2
; Sequence 2, Application US/10071976
; Publication No. US20030036178A1
; GENERAL INFORMATION:
; APPLICANT: Tsien et al., Roger Y. et al.
; TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT
; PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/071,976
FILING DATE: 05-Feb-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/465,142
FILING DATE: 1999-DEC-16
APPLICATION NUMBER: 08/974,737
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/056001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-071-976-2

Query Match 100.0%; Score 1270; DB 14; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.8e-119;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVDGDNVGHKFSVSGEGEDVTYKLTCLKFTCTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVDGDNVGHKFSVSGEGEDVTYKLTCLKFTCTTGKLPVWPPTL 60
Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180

Qy 181 YQOQNTPLDGPVLLPDNHYLSQTSALSQKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YQOQNTPLDGPVLLPDNHYLSQTSALSQKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 3

US-10-620-099-2
; Sequence 2, Application US/10620099
; Publication No. US20040014128A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: WACHTER, Rebekka
; APPLICANT: REMINGTON, James
; TITLE OF INVENTION: LONG WAVELENGTH ENGINEERED FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1250-5
; CURRENT APPLICATION NUMBER: US/10/620,099
; CURRENT FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US/09/575,847
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 08/974,737
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: US 08/911,825
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 08/706,408
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea

US-10-620-099-2

Query Match 100.0%; Score 1270; DB 15; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.8e-119;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVPIVLVDGDNVGHKFSVSGEGEDVTYKLTCLKFTCTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVDGDNVGHKFSVSGEGEDVTYKLTCLKFTCTTGKLPVWPPTL 60
Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YQOQNTPLDGPVLLPDNHYLSQTSALSQKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YQOQNTPLDGPVLLPDNHYLSQTSALSQKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 4

US-10-924-232-2
; Sequence 2, Application US/10924232
; Publication No. US20050079525A1
; GENERAL INFORMATION:
; APPLICANT: Tsien et al., Roger Y. et al.
; TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT
; PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/924,232
FILING DATE: 23-Aug-2004
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/071,976
FILING DATE: 05-Feb-2002
APPLICATION NUMBER: 09/465,142
FILING DATE: 1999-DEC-16
APPLICATION NUMBER: 08/974,737
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/056001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-924-232-2

Query Match 100.0%; Score 1270; DB 17; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.8e-119;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVDGVDVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVDGVDVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVWPPTL 60
Qy 61 VTTFSYGVQCFSRYPDMHKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDMHKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy 121 NRLEKIDFKEDGNILGHKLEYNYNHNVYIMADKQNGIKVNFKIRHNIEDGSQLAD 180
Db 121 NRLEKIDFKEDGNILGHKLEYNYNHNVYIMADKQNGIKVNFKIRHNIEDGSQLAD 180
Qy 181 YVQONTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 YVQONTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 5
US-09-884-681-2
Sequence 2, Application US/09884681
Patent No. US20020061546A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
Cubitt, Andrew B.
TITLE OF INVENTION: Assays for Protein Kinases Using
Correspondence Address:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/884,681
FILING DATE: 19-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/679,865
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John S.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 023072-069000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-884-681-2

Query Match 98.1%; Score 1246; DB 9; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.7e-117;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVDGVDVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVDGVDVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVWPPTL 60
Qy 61 VTTFSYGVQCFSRYPDMHKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDMHKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy 121 NRLEKIDFKEDGNILGHKLEYNYNHNVYIMADKQNGIKVNFKIRHNIEDGSQLAD 180
Db 121 NRLEKIDFKEDGNILGHKLEYNYNHNVYIMADKQNGIKVNFKIRHNIEDGSQLAD 180
Qy 181 YVQONTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 YVQONTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 6
US-09-967-301-2
Sequence 2, Application US/09967301
Publication No. US20030175859A1
GENERAL INFORMATION:
APPLICANT: Stubbs, Simon L.
APPLICANT: Jones, Anne B.
APPLICANT: Michael, Nigel P.
APPLICANT: Thomas, Nicholas
TITLE OF INVENTION: Fluorescent Proteins
FILE REFERENCE: PA0111
CURRENT APPLICATION NUMBER: US/09/967,301
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: GB 0109858.1
PRIOR FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea victoria
US-09-967-301-2

Query Match 98.1%; Score 1246; DB 10; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.7e-117;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVDGVDVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVDGVDVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVWPPTL 60

Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRTELKGDIFKEDGNILGHKLEYNYNHSHVYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRTELKGDIFKEDGNILGHKLEYNYNHSHVYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YQOQNTPIIDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 YQOQNTPIIDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 7
US-10-024-686-2
; Sequence 2, Application US/10024686
; Publication No. US20020123113A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; Heim, Roger
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/024,686
; FILING DATE: 17-Dec-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/057,995
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/727,452
; FILING DATE: 10-OCT-1996
; APPLICATION NUMBER: US95/14692
; FILING DATE: 13-NOV-1995
; APPLICATION NUMBER: 08/337,915
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/032002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-024-686-2

Query Match 98.1%; Score 1246; DB 13; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.7e-117;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVPILVELDGVNGHKFVS GEGEDVTYKGLTKFKICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPIVLVDGVNGHKFVS GEGEDATYKGLTKFKICTTGKLPVPWPTL 60
Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120

Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRTELKGDIFKEDGNILGHKLEYNYNHSHVYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRTELKGDIFKEDGNILGHKLEYNYNHSHVYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YQOQNTPIIDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 YQOQNTPIIDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 8
US-10-057-505-2
; Sequence 2, Application US/10057505
; Publication No. US20020164674A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: AURORA BIOSCIENCES CORPORATION
; APPLICANT: TSIEH, Roger
; APPLICANT: HEIM, Roger
; APPLICANT: CUBITT, Andrew
; TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS
; FILE REFERENCE: REGEN1260-3
; CURRENT APPLICATION NUMBER: US/10/057,505
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 08/792,553
; PRIOR FILING DATE: 1997-01-31
; PRIOR APPLICATION NUMBER: US 09/396,003
; PRIOR FILING DATE: 1999-09-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-057-505-2

Query Match 98.1%; Score 1246; DB 13; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.7e-117;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVPILVELDGVNGHKFVS GEGEDVTYKGLTKFKICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPIVLVDGVNGHKFVS GEGEDATYKGLTKFKICTTGKLPVPWPTL 60
Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRTELKGDIFKEDGNILGHKLEYNYNHSHVYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRTELKGDIFKEDGNILGHKLEYNYNHSHVYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YQOQNTPIIDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 YQOQNTPIIDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 9
US-10-293-580-2
; Sequence 2, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AU01270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/10/293,580
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US/09/129,192
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea
US-10-293-580-2

Query Match      98.1%; Score 1246; DB 14; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.7e-117;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVDGVDNGHKFSVSGEGEDVTYKGLTLKFKICTTGKLPVWPPTL 60
    |||||
Db 1 MSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGEDATYKGLTLKFKICTTGKLPVWPPTL 60
    |||||

Qy 61 VTTFSYGVQCFSRYPDMKRDHDFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
    |||||
Db 61 VTTFSYGVQCFSRYPDMKRDHDFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
    |||||

Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
    |||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
    |||||

Qy 181 YQQNTPIGDPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
    :|||
Db 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
    :|||

RESULT 10
US-10-293-580-74
; Sequence 74, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/10/293,580
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US/09/129,192
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea green fluorescent protein phosphorylation mutant
US-10-293-580-74

Query Match      98.1%; Score 1246; DB 14; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.7e-117;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVDGVDNGHKFSVSGEGEDVTYKGLTLKFKICTTGKLPVWPPTL 60
    |||||
Db 1 MSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGEDATYKGLTLKFKICTTGKLPVWPPTL 60
    |||||

Qy 61 VTTFSYGVQCFSRYPDMKRDHDFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
    |||||
Db 61 VTTFSYGVQCFSRYPDMKRDHDFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
    |||||

Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
    |||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
    |||||

Qy 181 YQQNTPIGDPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
    :|||
Db 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
    :|||

RESULT 11
US-10-457-982-2
; Sequence 2, Application US/10457982
```

```
; Publication No. US20030212265A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Liopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/10/457,982
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: US/09/602,641
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/172,063
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-457-982-2

Query Match      98.1%; Score 1246; DB 15; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.7e-117;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVDGVDNGHKFSVSGEGEDVTYKGLTLKFKICTTGKLPVWPPTL 60
    |||||
Db 1 MSKGEELFTGVVPIVLVDGVDNGHKFSVSGEGEDATYKGLTLKFKICTTGKLPVWPPTL 60
    |||||

Qy 61 VTTFSYGVQCFSRYPDMKRDHDFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
    |||||
Db 61 VTTFSYGVQCFSRYPDMKRDHDFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
    |||||

Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
    |||||
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
    |||||

Qy 181 YQQNTPIGDPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
    :|||
Db 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
    :|||

RESULT 12
US-10-724-178-2
; Sequence 2, Application US/10724178
; Publication No. US20040137528A1
; GENERAL INFORMATION:
; APPLICANT: Odyssey Thera, Inc.
; APPLICANT: Michnick, Stephen
; APPLICANT: Macdonald, Marnie
; APPLICANT: Lamerdin, Jane
; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT
; TITLE OF INVENTION: COMPLEMENTATION ASSAYS
; FILE REFERENCE: ODDY007
; CURRENT APPLICATION NUMBER: US/10/724,178
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US 60/461,133
; PRIOR FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-724-178-2

Query Match      98.1%; Score 1246; DB 16; Length 238;
Best Local Similarity 97.9%; Pred. No. 4.7e-117;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```


Qy	121	NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQKNGIKVNFKIRHNIEDGSVOLAD	180
Db	121	NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQKNGIKVNFKIRHNIEDGSVOLAD	180
Qy	181	YQONTPILDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDLYK	238
Db	181	YQONTPILDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDLYK	238

Search completed: April 25, 2005, 15:59:51
Job time : 138 secs

THIS PAGE BLANK (U.S.P.C.)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 15:37:46 ; Search time 43 Seconds
(without alignments)
413.174 Million cell updates/sec

Title: US-10-620-099-2

Perfect score: 1270

Sequence: 1 MSKGEELFTAVVPILVELDG.....VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1270	100.0	238	1	US-08-911-825-2
2	1270	100.0	238	2	US-08-911-825-2
3	1270	100.0	238	3	US-08-911-825-2
4	1270	100.0	238	4	US-08-911-825-2
5	1270	100.0	238	5	US-08-911-825-2
6	1270	100.0	238	6	US-08-911-825-2
7	1246	98.1	238	1	US-08-753-143-2
8	1246	98.1	238	2	US-08-679-865-2
9	1246	98.1	238	2	US-08-680-876-2
10	1246	98.1	238	2	US-08-792-553-2
11	1246	98.1	238	3	US-08-753-144-2
12	1246	98.1	238	3	US-08-094-359-2
13	1246	98.1	238	3	US-08-172-063-2
14	1246	98.1	238	3	US-08-263-975-2
15	1246	98.1	238	3	US-08-727-452-2
16	1246	98.1	238	4	US-08-418-785-1
17	1246	98.1	238	4	US-08-129-192C-2
18	1246	98.1	238	4	US-08-129-192C-74
19	1246	98.1	238	4	US-08-602-641-2
20	1246	98.1	238	4	US-08-704-463-2
21	1246	98.1	238	4	US-10-024-686A-2
22	1246	98.1	238	4	US-08-594-575C-2
23	1242	97.8	238	1	US-08-337-915A-2
24	1242	97.8	238	3	US-08-121-539-1
25	1242	97.8	238	4	US-08-214-909-2
26	1242	97.8	238	4	US-08-479-645A-10
27	1242	97.8	238	4	US-08-479-645A-159

28	1242	97.8	238	4	US-09-472-065A-4	Sequence 4, Appli
29	1242	97.8	238	4	US-09-920-922-4	Sequence 4, Appli
30	1242	97.8	238	5	PCT-US95-14692-2	Sequence 2, Appli
31	1241	97.7	238	4	US-09-023-946B-35	Sequence 35, Appli
32	1241	97.7	239	3	US-08-646-538-2	Sequence 2, Appli
33	1241	97.7	239	3	US-09-503-222-2	Sequence 2, Appli
34	1239	97.6	238	4	US-09-472-065A-2	Sequence 2, Appli
35	1239	97.6	243	4	US-09-479-645A-94	Sequence 94, Appli
36	1239	97.6	243	4	US-09-479-645A-96	Sequence 96, Appli
37	1239	97.6	243	4	US-09-479-645A-98	Sequence 98, Appli
38	1239	97.6	243	4	US-09-479-645A-100	Sequence 100, App
39	1239	97.6	243	4	US-09-479-645A-102	Sequence 102, App
40	1239	97.6	243	4	US-09-479-645A-104	Sequence 104, App
41	1239	97.6	243	4	US-09-479-645A-110	Sequence 110, App
42	1239	97.6	1070	4	US-09-091-042A-2	Sequence 2, Appli
43	1239	97.6	1452	3	US-09-127-227-2	Sequence 2, Appli
44	1238	97.5	243	4	US-09-479-645A-88	Sequence 88, Appli
45	1238	97.5	243	4	US-09-479-645A-90	Sequence 90, Appli

ALIGNMENTS

RESULT 1
US-08-911-825-2
; Sequence 2, Application US/08911825
; Patent No. 6054321
; GENERAL INFORMATION:
; APPLICANT: Tsien et al., Roger Y. et al.
; TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,825
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/056001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-911-825-2

Query Match 100.0%; Score 1270; DB 3; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.1e-130;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVPILVELDGVNKHKFSVSGEGDVTYKGLTKFKICTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPILVELDGVNKHKFSVSGEGDVTYKGLTKFKICTTGKLPVWPPTL 60
Qy 61 VTTFSYGVQCFSRYPDHMKSHDFKFSAMPEGYVQORTIFFKDDGNKYKRAEVKPEGDTLV 120

Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
QY 121 NRLEKIGIDFKEDGNILGHKLEYNYNHSHNYYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRLEKIGIDFKEDGNILGHKLEYNYNHSHNYYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
QY 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDHVMVLEFVTAAGITHGMDLYK 238
Db 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDHVMVLEFVTAAGITHGMDLYK 238

RESULT 2

US-08-974-737-2
; Sequence 2, Application US/08974737
; Patent No. 6077707
; GENERAL INFORMATION:
; APPLICANT: Tsien et al., Roger Y. et al.
; TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,737
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,825
; FILING DATE: 15-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/056001
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-974-737-2

Query Match 100.0%; Score 1270; DB 3; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.1e-130;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGEDVTYKGLTLKFTCTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGEDVTYKGLTLKFTCTTGKLPVWPPTL 60
QY 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
QY 121 NRLEKIGIDFKEDGNILGHKLEYNYNHSHNYYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRLEKIGIDFKEDGNILGHKLEYNYNHSHNYYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
QY 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDHVMVLEFVTAAGITHGMDLYK 238
Db 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDHVMVLEFVTAAGITHGMDLYK 238

RESULT 3

US-08-706-408-2
; Sequence 2, Application US/08706408
; Patent No. 6124128
; GENERAL INFORMATION:
; APPLICANT: Tsien et al., Roger Y.
; TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,408
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,050
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/029001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-706-408-2

Query Match 100.0%; Score 1270; DB 3; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.1e-130;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGEDVTYKGLTLKFTCTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGEDVTYKGLTLKFTCTTGKLPVWPPTL 60
QY 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
QY 121 NRLEKIGIDFKEDGNILGHKLEYNYNHSHNYYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRLEKIGIDFKEDGNILGHKLEYNYNHSHNYYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
QY 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDHVMVLEFVTAAGITHGMDLYK 238
Db 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDHVMVLEFVTAAGITHGMDLYK 238

RESULT 4

US-09-465-142-2
; Sequence 2, Application US/09465142
; Patent No. 6403374
; GENERAL INFORMATION:
; APPLICANT: Tsien et al., Roger Y. et al.
; TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT

;; TITLE OF INVENTION: PROTEINS
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson P.C.
;; STREET: 4225 Executive Square, Suite 1400
;; CITY: La Jolla
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 92037
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/465,142
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/974,737
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haile, Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: 07257/056001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619/678-5099
;; TELEFAX: 619/678-5099
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 238 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-465-142-2

Query Match 100.0%; Score 1270; DB 4; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.1e-130;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPLVLDGVDVNGHKFSVSGEGDVTYKGLTLKFKICTTGKLPVPMPTL 60
Db 1 MSKGEELFTAVVPLVLDGVDVNGHKFSVSGEGDVTYKGLTLKFKICTTGKLPVPMPTL 60

Qy 61 VTTFSYGVQCFSRYPDHHKRDHFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHHKRDHFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180

Qy 181 YQONTPILOGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YQONTPILOGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 5
US-09-575-847-2
; Sequence 2, Application US/09575847
; Patent No. 6593135
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: WACHTER, Rebekka
; APPLICANT: REMINGTON, James
; TITLE OF INVENTION: LONG WAVELENGTH ENGINEERED FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1250-5
; CURRENT APPLICATION NUMBER: US/09/575,847
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 08/974,737
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: US 08/911,825
; PRIOR FILING DATE: 1997-08-15

;; PRIOR APPLICATION NUMBER: US 08/706,408
;; PRIOR FILING DATE: 1996-08-30
;; NUMBER OF SEQ ID NOS: 20
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 2
;; LENGTH: 238
;; TYPE: PRT
;; ORGANISM: Aequorea
;; US-09-575-847-2

Query Match 100.0%; Score 1270; DB 4; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.1e-130;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPLVLDGVDVNGHKFSVSGEGDVTYKGLTLKFKICTTGKLPVPMPTL 60
Db 1 MSKGEELFTAVVPLVLDGVDVNGHKFSVSGEGDVTYKGLTLKFKICTTGKLPVPMPTL 60

Qy 61 VTTFSYGVQCFSRYPDHHKRDHFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHHKRDHFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180

Qy 181 YQONTPILOGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YQONTPILOGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 6
US-10-071-976-2
; Sequence 2, Application US/10071976
; Patent No. 6780975
; GENERAL INFORMATION:
; APPLICANT: Tsien et al., Roger Y. et al.
; TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT
; PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/071,976
; FILING DATE: 05-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/465,142
; FILING DATE: 1999-DEC-16
; APPLICATION NUMBER: 08/974,737
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/056001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-071-976-2

Query Match      100.0%; Score 1270; DB 4; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.1e-130;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVELDGVNGHKFVSVEGEGDVTYGLKTLKFCITTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVELDGVNGHKFVSVEGEGDVTYGLKTLKFCITTTGKLPVWPPTL 60

Qy 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNHSHVYIMADKQKGIKYNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNHSHVYIMADKQKGIKYNFKIRHNIEDGSVOLAD 180

Qy 181 YVQONTPIIDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YVQONTPIIDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 7
US-08-753-143-2
; Sequence 2, Application US/08753143A
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; FILE REFERENCE: 07257/032003
; CURRENT APPLICATION NUMBER: US/08/753.143A
; CURRENT FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-08-753-143-2

Query Match      98.1%; Score 1246; DB 1; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVELDGVNGHKFVSVEGEGDVTYGLKTLKFCITTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVELDGVNGHKFVSVEGEGDVTYGLKTLKFCITTTGKLPVWPPTL 60

Qy 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNHSHVYIMADKQKGIKYNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNHSHVYIMADKQKGIKYNFKIRHNIEDGSVOLAD 180

Qy 181 YVQONTPIIDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YVQONTPIIDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 8
US-08-679-865-2
; Sequence 2, Application US/08679865
; Patent No. 5912137
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/679,865
; FILING DATE: 16-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John S.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 023072-069000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-679-865-2

Query Match      98.1%; Score 1246; DB 2; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVELDGVNGHKFVSVEGEGDVTYGLKTLKFCITTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVELDGVNGHKFVSVEGEGDVTYGLKTLKFCITTTGKLPVWPPTL 60

Qy 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120

Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNHSHVYIMADKQKGIKYNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNHSHVYIMADKQKGIKYNFKIRHNIEDGSVOLAD 180

Qy 181 YVQONTPIIDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YVQONTPIIDGVPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 9
US-08-680-876-2
; Sequence 2, Application US/08680876
; Patent No. 5925558
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,876
FILING DATE: 16-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John S.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 023072-069200
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-680-876-2

Query Match 98.1%; Score 1246; DB 2; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYKGLTLKFICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVILVELDGVNGHKFSVSGEGDVTYKGLTLKFICTTGKLPVPWPTL 60
Qy 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
Qy 181 YYQNTPTILDGVPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 HYQNTPTIGDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 10
US-08-792-553-2
Sequence 2, Application US/08792553
Patent No. 5981200
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
TITLE OF INVENTION: Tandem Fluorescent Protein Constructs
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & RICHARDSON P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/792,553
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lisa A. Haile, Ph.D.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/041001/UC 96-160-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070

TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-792-553-2
Query Match 98.1%; Score 1246; DB 2; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYKGLTLKFICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVILVELDGVNGHKFSVSGEGDVTYKGLTLKFICTTGKLPVPWPTL 60
Qy 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
Qy 181 YYQNTPTILDGVPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 HYQNTPTIGDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
RESULT 11
US-08-753-144-2
Sequence 2, Application US/08753144
Patent No. 6066476
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/753,144
FILING DATE: 20-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/727,452
FILING DATE: 10-OCT-1996
APPLICATION NUMBER: US95/14692
FILING DATE: 13-NOV-1995
APPLICATION NUMBER: 08/337,915
FILING DATE: 10-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/032002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear

```
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-753-144-2

Query Match          98.1%; Score 1246; DB 3; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGDVYTKLTLKFKICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPILVELDGVNGHKFSVSGEGDATYKGLTLKFKICTTGKLPVWPPTL 60
Qy 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YQONTPIIDGPVLLPDNHYLSQTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIIDGPVLLPDNHYLSQTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 12
US-09-094-359-2
; Sequence 2, Application US/09094359
; Patent No. 6140132
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Liopis, Juan
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/067001
; CURRENT APPLICATION NUMBER: US/09/094,359
; CURRENT FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-094-359-2

Query Match          98.1%; Score 1246; DB 3; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGDVYTKLTLKFKICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPILVELDGVNGHKFSVSGEGDATYKGLTLKFKICTTGKLPVWPPTL 60
Qy 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YQONTPIIDGPVLLPDNHYLSQTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIIDGPVLLPDNHYLSQTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 13
US-09-172-063-2
; Sequence 2, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
```

```
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Liopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
; EARLIER FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-172-063-2

Query Match          98.1%; Score 1246; DB 3; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGDVYTKLTLKFKICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPILVELDGVNGHKFSVSGEGDATYKGLTLKFKICTTGKLPVWPPTL 60
Qy 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YQONTPIIDGPVLLPDNHYLSQTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIIDGPVLLPDNHYLSQTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 14
US-09-263-975-2
; Sequence 2, Application US/09263975
; Patent No. 6248550
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/263,975
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/679,865
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John S.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 023072-069000
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-263-975-2

Query Match 98.1%; Score 1246; DB 3; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MSKGEELFTAVPILVELDGVNGHKFVSVEGEGDVTYKLTCLKFTCTTGKLPVPWPTL 60
DB 1 MSKGEELFTAVPILVELDGVNGHKFVSVEGEGDVTYKLTCLKFTCTTGKLPVPWPTL 60
QY 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
QY 181 YQONTPILDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
DB 181 YQONTPILDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 15
US-08-727-452-2
Sequence 2, Application US/08727452A
Patent No. 6319669
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Heim, Roger
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR DETECTION OF ANALYTES
FILE REFERENCE: 07257/032001
CURRENT APPLICATION NUMBER: US/08/727,452A
CURRENT FILING DATE: 1996-03-20
EARLIER APPLICATION NUMBER: PCT/US95/14692
EARLIER FILING DATE: 1995-11-13
EARLIER APPLICATION NUMBER: US 07/337,915
EARLIER FILING DATE: 1994-11-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea victoria
US-08-727-452-2

Query Match 98.1%; Score 1246; DB 3; Length 238;
Best Local Similarity 97.9%; Pred. No. 1.3e-127;
Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MSKGEELFTAVPILVELDGVNGHKFVSVEGEGDVTYKLTCLKFTCTTGKLPVPWPTL 60
DB 1 MSKGEELFTAVPILVELDGVNGHKFVSVEGEGDVTYKLTCLKFTCTTGKLPVPWPTL 60
QY 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
QY 181 YQONTPILDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
DB 181 YQONTPILDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

Db 181 YQONTPILDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Search completed: April 25, 2005, 15:48:17
Job time : 45 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 15:36:47 ; Search time 40 Seconds
(without alignments)
572.490 Million cell updates/sec

Title: US-10-620-099-2

Perfect score: 1270

Sequence: 1 MSKGELFTAVVPILVELDG.....VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1236	97.3	238	1 JQ1514	green-fluorescent
2	105.5	8.3	785	2 H72228	hypothetical prote
3	100.5	7.9	887	2 E82590	leucyl-tRNA synth
4	99	7.8	336	2 C64468	hypothetical prote
5	92.5	7.3	2573	2 D71614	hypothetical prote
6	90	7.1	471	2 T27856	hypothetical prote
7	90	7.1	797	2 JC4078	protective surface
8	90	7.1	808	2 F64102	protective surface
9	89.5	7.0	393	2 C64613	conserved hypotet
10	88.5	7.0	312	2 C81710	thioredoxin reduct
11	88.5	7.0	425	2 C97354	hypothetical prote
12	87.5	6.9	861	2 H64102	leucine-tRNA ligas
13	87	6.9	822	2 C71633	ATP-dependent nucl
14	87	6.9	2334	2 S32920	cell wall-associat
15	86	6.8	357	2 G81355	tRNA (uracil-5-)-m
16	85.5	6.7	752	1 KXRTC1	proprotein convert
17	85.5	6.7	836	1 JDUVLD	DNA-directed DNA p
18	85.5	6.7	889	2 JC5576	inter-alpha-trypsi
19	84.5	6.7	351	2 B71556	probable thioredox
20	84.5	6.7	860	2 AC0582	leucyl-tRNA synth
21	83	6.5	281	2 AD2052	hypothetical prote
22	83	6.5	533	2 S47271	ANAK-related prote
23	83	6.5	1092	2 S42798	fibronectin-bindin
24	83	6.5	3472	2 T31308	hypothetical 367K
25	82.5	6.5	613	2 A99552	oligopeptidase
26	82.5	6.5	1134	2 A60234	IGA Fc receptor pr
27	82.5	6.5	1164	1 FC02AG	IGA Fc receptor pr
28	82	6.5	207	2 C70483	hypothetical prote
29	82	6.5	653	1 A11535	fructose-bisphosph

ALIGNMENTS

RESULT 1

JQ1514

green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)

C:Species: Aequorea victoria

C:Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 09-Jul-2004

C:Accession: JS0692; JQ1514; F00335; S48693; S51330; S51331

R:Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.

Gene 111, 229-233, 1992

A:Title: Primary structure of the Aequorea victoria green-fluorescent protein.

A:Reference number: JQ1514; MUID:92175527; PMID:1347277

A:Accession: JS0692

A:Molecule type: DNA

A:Residues: 1-107, 'S', 109-238 <PRA1>

A:Cross-references: UNIPROT:P42212; UNIPROT:Q17106; UNIPROT:Q17105; GB:M62654; NID:g15

A:Accession: JQ1514

A:Molecule type: mRNA

A:Residues: 1-99, 'P', 101-140, 'L', 142-218, 'V', 220-238 <PRA2>

A:Cross-references: GB:M62653; NID:g155660; PIDN:AAA27721.1; PID:g155661

A:Accession: F00335

A:Molecule type: protein

A:Residues: 46-64; 74-122; 132-151; 154-183; 185-200 <PRA3>

R:Inouye, S.; Tsuji, F.I.

FEBS Lett. 351, 211-214, 1994

A:Title: Evidence for redox forms of the Aequorea green fluorescent protein.

A:Reference number: S48693; MUID:94364470; PMID:8082767

A:Accession: S48693

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-24, 'Q', 26-156, 'P', 158-171, 'K', 173-238 <INO>

A:Cross-references: GB:I29345; NID:g606383; PIDN:AAA58246.1; PID:g606384

R:Watkins, J.N.; Campbell, A.K.

submitted to the EMBL Data Library, January 1995

A:Reference number: S51330

A:Accession: S51330

A:Molecule type: mRNA

A:Residues: 1-13, 'V', 15-24, 'Q', 26-44, 'N', 46-153, 'G', 155-156, 'P', 158-171, 'K', 173-227, 'R', 'Q'

A:Cross-references: EMBL:X83959; NID:g634008; PIDN:CAA58789.1; PID:g634009

A:Experimental source: clone gfp1

A:Accession: S51331

A:Molecule type: mRNA

A:Residues: 1-24, 'Q', 26-29, 'R', 31-83, 'L', 85-153, 'G', 155-156, 'P', 158-171, 'K', 173-208, 'Q', 'R', 'Q'

A:Cross-references: EMBL:X83960; NID:g634010; PIDN:CAA58790.1; PID:g634011

A:Experimental source: clone gfp2

R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.

submitted to the Brookhaven Protein Data Bank, August 1996

A:Reference number: A65692; PDB:1GFL

A:Contents: annotation: X-ray crystallography, 1.9 angstroms, residues 'A', 2-79, 'R', '81'

A:Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli

R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.

Nat. Biotechnol. 14, 1246-1251, 1996

A:Title: The molecular structure of green fluorescent protein.

A:Reference number: A58953; MUID:98294543; PMID:9631087

A;Contents: annotation; X-ray crystallography, 1.9 angstroms
C;Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJFNV) emitting
C;Comment: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr-
C;Genetics:

A;Gene: GFP
A;Introns: 69/3; 167/3
C;Superfamily: green-fluorescent protein
C;Keywords: chromoprotein; luminescence
F;65-67/Cross-link: 5-imidazolone (Ser-Gly) #status experimental
F;66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match 97.3%; Score 1236; DB 1; Length 238;
Best Local Similarity 96.2%; Pred. No. 1.2e-95;
Matches 229; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVVPIVLVELDGVNKHFSVSGEGDVTYVKLTIKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVPIVLVELDGVNKHFSVSGEGDVTYVKLTIKFICTTGKLPVWPPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHPFKSAMPEGYVQQTIFPKDGNKYKTRAEVKFEGDITLV 120
DB 61 VTTFSGVQCFSRYPDHMKRHPFKSAMPEGYVQQTIFPKDGNKYKTRAEVKFEGDITLV 120
QY 121 NRTELKIDPKEDGNILGHKLEYNVSHNVYINADKQKGIKVNFKIRHNIEDGSVOLAD 180
DB 121 NRTELKIDPKEDGNILGHKLEYNVSHNVYINADKQKGIKVNFKIRHNIEDGSVOLAD 180
QY 181 YVQONTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDLYK 238
DB 181 HYQONTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDLYK 238

RESULT 2
H72228
hypothetical protein TM1624 - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: H72228
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
C.M.

Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: H72228
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-785 <ARN>
A;Cross-references: UNIPROT:Q9X1V9; GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AAD3669

A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1624

Query Match 8.3%; Score 105.5; DB 2; Length 785;
Best Local Similarity 23.9%; Pred. No. 0.84; Indels 47; Gaps 7;
Matches 43; Conservative 25; Mismatches 65;

QY 21 DVNKGH-----KFSVSGEGEDVTYVKLTIKFICTTGKLPVWPPTLVTFSGVQCFS 72
DB 5 DLNGFVSRVDRNEGRFSGEIVPG-VVQADLVK-----GLLPHPYVGM----- 46
QY 73 RYPDHMKRHPFKSAMPEGYVQQTIFPKDGNKYKTRAEVKFEGDITLVNRIELKGDIFKE 132
DB 47 -----NEDLFKEIDREWIYEREFKEDVKEGERVLDVFEVDITSLVDYLVNGVY-- 97
QY 133 DGNILGHKLEYNVSHNVYINADKQKGIKVNFKIRHNIEDGSVOLADYVQONTPIIDGP 192
DB 98 -GSTEDMFIERFDVTNL-----KEKNHLKYIK-----SPIRVPKTLEQNYGVLGPP 145

RESULT 3
E82590
leucyl-tRNA synthetase XF2176 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: E82590
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ
Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below

A;Accession: E82590
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-887 <SIM>
A;Cross-references: GB:AE004031; GB:AE003849; NID:g9107309; PIDN:AAF84975.1; GSPDB:GN00
A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir,
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;

A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2176
C;Superfamily: leucine-tRNA ligase

Query Match 7.9%; Score 100.5; DB 2; Length 887;
Best Local Similarity 23.2%; Pred. No. 2.6;
Matches 46; Conservative 28; Mismatches 67; Indels 57; Gaps 10;

QY 49 TTCKLPVPWPTLVTTTFSGVQCFSRYPDHMKRHPFKSAMPEGYVQQTIFPKDGNV-- 106
DB 329 TNEQLPV-WVAFVLMVAYGTGAVMVGHDQDQEF--ANKYGLPFRQVIALKEPKNQDE 385
QY 107 -----KTRAEVKFEGDITLVNRIELKGDIFDKDGNILGHKLEYNVSHNVYI 152

DB 386 STWEPDVRWDYADKTR---EFE---LINSAEFDGLDQDAPEVLAERFE----- 429
QY 153 MADKQKNG-IKNFKIRHNIEDGSVOLADYVQONTPI-----DGPVLLPDN 198
DB 430 ---RQGRQQRVNYRLR----DWGVSQRQYWGCPFIVYICPTCGAVPVPEDQLPVILPEN 482

QY 199 -HYLSTQSALSKDPNEKR 215
DB 483 VAFSGTGSPIKTDPEWRK 500

RESULT 4
C64468
hypothetical protein MJ1348 - Methanococcus jannaschii

C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: C64468

R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschi
A;Reference number: A64300; MUID:96337999; PMID:8688087

A;Accession: C64468
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

A;Residues: 1-336 <BUL>
A;Cross-references: UNIPROT:Q58743; GB:U67574; GB:L77117; NID:g1591978; PIDN:AAB99360.1
C;Genetics:
A;Map position: FOR1295121-1296131

QY 215 RDHWLLEFVTAAGITHG 232
Db 615 RDHLWVVSASAKAGYANG 632

RESULT 8

F64102
conserved hypothetical protein D-15 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Sep-1998
C;Accession: F64102
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: F64102
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-808 <TIGR>
A;Cross-references: GB:L42023; TIGR:HI0917
C;Superfamily: protective surface antigen D-15
C;Keywords: surface antigen

Query Match 7.1%; Score 90; DB 2; Length 808;
Best Local Similarity 20.2%; Pred. No. 17;
Matches 52; Conservative 36; Mismatches 96; Indels 74; Gaps 12;

QY 22 VNGHK-----FVSSEGEQDVYTKLTKFICTTKLPVWPVLVTTFSYGVQCFSRYP 75
Db 415 INGSNDVDVYKVKRNTGSINFG-----ICYGTEGSIQY 451

QY 76 DHMKRHF-----FKSAMPEGYVQ-----RTIFFKDDGNKTRA 110
Db 452 ASVKQDNFLGTGAASVIAGTNDYGSVNLGYTEPYFTKDGSLGGNVFFENYDNSKDT 511

QY 111 EVKFEQDVLNRIELKIDFKDGNI---LGH-----KLEYNNSHNVYIMADKQK 158
Db 512 SSNYKRTYGSNVTL-GFPVNNNSYVGLHTYKISNPALEYN---RNLIQSMKFG 567

QY 159 NGIKVN-FKIRHNIEQSVOLADYQYQNTPI-LDGPVLPLP--DNHYLSTQSALSADPNK 214
Db 568 NGIKTNDFDSFGWYNNLSNRGYFTPKGVKASLGSRVTTPGSDNKYKLSADYQGFYPLD 627

QY 215 RDHWLLEFVTAAGITHG 232
Db 628 RDHLWVVSASAKAGYANG 645

RESULT 9

C64613
conserved hypothetical protein HP0747 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: C64613
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen-
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: C64613
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-393 <TOM>
A;Cross-references: UNIPROT:O25443; GB:AE000587; GB:AE000511; NID:g2313869; PIDN:AAD0779

Query Match 7.0%; Score 89.5; DB 2; Length 393;
Best Local Similarity 20.8%; Pred. No. 7.3;

Matches 51; Conservative 42; Mismatches 77; Indels 75; Gaps 15;
QY 34 EGDVYTKLTK-----FICTGKLPVWPVLVTTFSYGVQCFSRYPDHMKRH-- 81
Db 174 QGD---GRLVLESMNHRCEKIFV---HFPVPW-----NEKKHRR 207

QY 82 ----DFFKSAM---PEGVVOQT---IFFKDD-----GNYKTRAEVKFEGD-TLVNRIE 124
Db 208 VLSEKFLNALRVLPKRGFLERLTDSDLSYFESDLKALKNFQCEIEIKKNAQIPVVSUYE 267
QY 125 LKGIQFKED-GNILGHKLEYN---YNSHNVYIMADK---QKNGIKVNFKIRHNIEDG-SV 176
Db 268 ARWNKLLKDDIYDLRIYSLEWNETPPDNHAFDFSEDTITISKKSVMGVTILTKKIIQSGYFV 327
QY 177 QIADYVQYQNTPII-----DGP-----VLLPDNHYLSTQSALSADPNKRDHWLLEFV 224
Db 328 HVCNIIYENKGDPLVELSMGDFWPVRLFLVLLTENQIFLYLNKSPKLTNNHKAHLQLLNTL 387
QY 225 TAAGI 229
Db 388 SOKGI 392

RESULT 10

C81710
thioredoxin reductase frameshift TC0375 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: C81710
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81710
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <TET>
A;Cross-references: UNIPROT:Q9PKT7; GB:AE002304; GB:AE002160; NID:g7190409; PIDN:AAF392
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0375
C;Superfamily: thioredoxin reductase; thioredoxin reductase homology

Query Match 7.0%; Score 88.5; DB 2; Length 312;
Best Local Similarity 21.1%; Pred. No. 6.6;
Matches 48; Conservative 29; Mismatches 81; Indels 69; Gaps 9;

QY 1 MSKGELEFVAVPILVELDGDYNGHKFSVSGEGEDVTYTKLTKFICTTKLPVWPPTL 60
Db 97 LKSGEDTFTCDACIIAT---GASAKRLSIPGAGDNEFWQKGVTAACAVC-DGASPI----- 147

QY 61 VTTFSYGVQCFSRYPDHMKRHDF---KSAMPEGYVQYQNTPIFFKDDGNYKTRAEVKFE 115
Db 148 -----FRDRDLFVIGGSDSALAEAMFLTR-----YGRKRVFVHR 181

QY 116 GDTL-----VNRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRH 169
Db 182 RDTLRASKAMVNAQ-----ANEKIVFLWNSEVVKILGDSLRSI-----DIFN 225

QY 170 NIEDGGSVOL-----ADYVQYQNTPILDGPVLPLPDNHYLSTQSALS 209
Db 226 NVEKTTVTWEAAGVFFAIGHQNPATLGGQLSDENGIIITEKSSR 272

RESULT 11

C97354
hypothetical protein CAC3702 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: C97354
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: C97354
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-425 <KUR>
A;Cross-references: UNIPROT:Q97CY8; GB:AE001437; PIDN:AAK01622.1; PID:g15026806; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3702

Query Match 7.0%; Score 88.5; DB 2; Length 425;
Best Local Similarity 24.3%; Pred. No. 9.9;
Matches 59; Conservative 29; Mismatches 82; Indels 73; Gaps 14;
Qy 20 GDVNGHFSVSGEGDVTYKLTLCFTCTGK-----LP-----VPW 57
Db 109 GFVNGKLLPASGEAIKD--FGKV-----LKNKGDKLGLTROVEMFGVGRVSGVGNIEF 162
Qy 58 PTLVTFYSYGQCFSRYPDHMKRHDFFKPSAMPEGVYQORTIFFKDDGNYKTR-----AEV 112
Db 163 SEVVKNFNGVGD-----KNFVKAVKSEDKVGNP---SGEGNYSTEHWYLYKE 211
Qy 113 KFEGD-----TLVNRIELKGIDPKE--DGNILGHKLEYN-----NSH- 148
Db 212 KYGEDNVVLTDEKSLNSVSEKLSFRKNGINDVYG-KTAVAYGENLDTVSMATSNHQ 270

Qy 149 -NVYIMADKQKNGIKVNFKIRHNIEDSGVQLADYQNTPIPDGPVLLPDNHYLSTQSAL 207
Db 271 YNPILTDEKFDKNMNYKILKEYGDESVAYSKY-EHTLTEDYIKTLPDEYF--TSKGL 327
Qy 208 SKD 210
Db 328 VKD 330

RESULT 12
H64102
leucine-tRNA ligase (EC 6.1.1.4) - Haemophilus influenzae (strain Rd KW20)
N;Alternate names: leucyl-tRNA synthetase
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: H64102
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: H64102
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-861 <TIGR>
A;Cross-references: UNIPROT:P43827; GB:U32774; GB:L42023; NID:g1573942; PIDN:AAC22581.1;
C;Genetics:
A;Gene: leus
C;Superfamily: leucine-tRNA ligase
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 6.9%; Score 87.5; DB 2; Length 861;
Best Local Similarity 24.3%; Pred. No. 30;
Matches 45; Conservative 25; Mismatches 68; Indels 47; Gaps 9;
Qy 49 TTGKLPVWPPLVTLTFYSYGQCFSRYPDHMKRHDFFKPSAMPEGVYQORTIFFKDDGNYKT 108
Db 314 TGDKLPVLPVTLTFYSYGQCFSRYPDHMKRHDFFKPSAMPEGVYQORTIFFKDDGNYKT 370
Qy 109 RAEVKEFGDITLVNRIELKGIDFKEDGILGHKLYNHNHNYIMADK-QKNGI---KVN 164
Db 371 TKQAFVHEHGKLVNSDEFGKNF--DGAFNG-----IADKLEKLVGKGRQVN 414

Qy 165 FKIRH-----NIEDGVSQVLADYQNTPIPDGPVLLPDNHYL-STQSAL 207
Db 415 YRLRDWGVSRQRYWGABIPMLTLENGDVPA-----PMEDLPILPDVMDGVKXSPI 467
Qy 208 SKDPN 212
Db 468 NADPN 472
RESULT 13
C71633
ATP-dependent nuclease chain A (adda) RP734 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: C71633
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark,
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: C71633
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-822 <AND>
A;Cross-references: UNIPROT:Q9ZCJ7; GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA151
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: adda; RP734

Query Match 6.9%; Score 87; DB 2; Length 822;
Best Local Similarity 25.2%; Pred. No. 31;
Matches 54; Conservative 28; Mismatches 54; Indels 78; Gaps 13;
Qy 57 WPTLVTFYSYGQCFSRYPDHMKRHDFFKPSAMPEGVYQORTIFFKDDGNYKTRAEVKPFG 116
Db 473 WPLLT-----NQKQKEFFWT-LPE-----DNKAKSAADLLI-- 503
Qy 117 DTLVN-----RIELKGI-----DFKEDGNILGHKLYNHNHNYIMADKQKNGIKVNF 165
Db 504 DKVWVFKEIKKSGILTSTASRISEKDFMLVKRD--KPSHN--LIKELSKAKLKEI 559
Qy 166 KIRHNIEDGVSQVLADYQNTPIPD-----GPVLLPDNHYLSTQSALS KDP----- 211
Db 560 SDRINL-----KENLPILDLTAAAKFVLLPDDDL--NLACLKSPILIGISEOKL 606
Qy 212 ----NEKRDHMLLEFFVTAAGITHGMD---ELYK 238
Db 607 YTLVKNKNDHTLWEVLSSHDKDIYHKLSIIEIYK 640

RESULT 14
S32920
cell wall-associated protein precursor wapa [similarity] - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: S32920; E69730; T47101
R;Foster, S.J.
Mol. Microbiol. 8, 299-310, 1993
A;Title: Molecular analysis of three major wall-associated proteins of Bacillus subtilis
protein.
A;Reference number: S32919; MUID:93302506; PMID:8316082
A;Accession: S32920
A;Molecule type: DNA
A;Residues: 1-2334 <FOS>
A;Cross-references: UNIPROT:Q07833; GB:L05634; NID:g304177; PIDN:AAA22883.1; PID:g30417
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berth
C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
leach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Serod, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69730
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2334 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:G2636442; PIDN:CAB15959.1; PID:G2636469
A;Experimental source: strain 168
R;Yoshida, K.; Sano, H.; Seki, S.; Oda, M.; Fujimura, M.; Fujita, Y.
Microbiology 141, 337-343, 1995
A>Title: Cloning and sequencing of a 29 kb region of the *Bacillus subtilis* genome containing
A;Reference number: Z24350; MUID:95219088; PMID:7704263
A;Accession: T47101
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2334 <YOS>
A;Cross-references: EMBL:D31856; NID:G603765; PIDN:BAA06656.1; PID:G603782
A;Experimental source: strain BGSC1A1
C;Genetics:
A;Gene: wapA; N170
C;Superfamily: cell wall-associated protein wapA

Query Match	6.9%;	Score 87;	DB 2;	Length 2334;		
Best Local Similarity	19.3%;	Pred. No. 1.2e+02;				
Matches	58;	Conservative 41;	Mismatches 105;	Indels 96;	Gaps 12;	
Qy	21	DVNGHKESVSGEGDVTY	--GKLTILKFTCTGKLPVPWPTLVTTSTSYGVQ	-----	69	
Db	1054	DENGHVTSITGPNKKVITYSYENDLLKKVTDTDG	-----TWTSYDYSBGRVLKQYS	1105		
Qy	70	-----CFSRYPDHMKRHDFFKSPAMEGVV	-----QQRTIF	-----FKDDG	104	
Db	1106	ANSTEAKPVFTFYQYSGHRLEKAINAKETTVYSYADAKKLTLLMTQNGRKRVQGYNEAG	1165			
Qy	105	N-----YKTRAEVKPEGTULNRIE	-----LKGIDFKEDGNILGHK	-----	140	
Db	1166	NPTQVDDAAGLKLTITTKYEGNNVEDVDPNVGTGKATSYQYDKGNVTSVKDAYGT	1225			
Qy	141	LEYNYSHN-VYIMADKQKN	-----GIKVPFKIRHNI	EDGVSQVLADYY	-----QQNTPI	188
Db	1226	ETVEYNNVDVTKMKDTTEGNVTDIAYDGLDAVSETDQSGKSSAAVYDKGNQIQSSKDL	1285			
Qy	189	LDGEVLLPDNNHYLSTQSAKSKDPNEKR	-----DHMVLLBEFTVTAAGT	ITHG	232	
Db	1286	SASTNIIUKDGSFPAOKSGWNLTKASDKRRKISVIADKSGVLSGSKALBVLGSOT	SAGTDHGG	1345		

RESULT 15

G81355
tRNA (uracil-5-)-methyltransferase (EC 2.1.1.135) Cj0831c [imported] - Campylobacter jejuni
C/Species: Campylobacter jejuni
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C/Accession: G81355
R/Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, R.P.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrett, T.
Nature 403, 665-668, 2000
A/Rtitle: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervirulence
A/Reference number: AB1250; UID:20150912; PMID:10688204
A/Accession: G81355
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-357 <PAR>
A/Cross-references: UNIPROT:Q9PP92; GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB7309
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/Gene: trmA; Cj0831c
C/Keywords: methyltransferase; S-adenosylmethionine

```

Query Match          6.8%; Score 86; DB 2; Length 357;
Best Local Similarity 21.8%; Pred. No. 13;
Matches 36; Conservative 23; Mismatches 44; Indels 62; Gaps 6
Qy 79 KRHDFKSAPEGVQVQRTTFFKDDGNVKRTAEVKF--EGDTLV----- 120
Db 14 EKHSFIKKYKFETTKDFKLFASDKDHYRTAEISFYHENDTFLYAMFDPKSKKYIIIEY 73
Qy 121 -----NRIELK-GIDPKEDGNILGHKLEYNN----- 146
Db 74 LDFADEXICAFMPRLLELYRQDNKLKEKLFGEVFLTTKQELSITLLYHKNIEDIKSNLEN 133
Qy 147 -SHNVYIMADKQNGIKVNFKIRH-----NIEDGSVOLADYQOON 185
Db 134 LSNILHILNARSKGKKLIIFKTENLROTLLNIQDRKI----FYEFN 174

Search completed: April 25, 2005, 15:47:37
Job time : 52 secs

```

Search completed: April 25, 2005, 15:47:37
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 15:25:20 ; Search time 180 Seconds
(without alignments)
677.082 Million cell updates/sec

Title: US-10-620-099-2
Perfect score: 1270
Sequence: 1 MSGEELFTAVPILVELDG.....VLLEFVTAAGTHGMDLYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1246	98.1	238	2 Q8GHE2	Q8Ghe2 azotobacter
2	1242	97.8	238	1 GFP_AEQVI	P42212 aequorea vi
3	1242	97.8	238	2 Q71RY9	Q71ry9 azotobacter
4	1241	97.7	238	2 Q8GHE4	Q8Ghe4 azomonas ag
5	1240	97.6	238	2 Q8GHE3	Q8Ghe3 azotobacter
6	1235	97.2	238	2 Q93125	Q93125 aequorea vi
7	1200	94.5	238	2 Q17105	Q17105 aequorea vi
8	1185	93.3	238	2 Q17106	Q17106 aequorea co
9	1172	92.3	238	2 Q6YX20	Q6Yx20 aequorea co
10	1079	85.0	238	2 Q8WP95	Q8Wp95 aequorea ma
11	1073	84.5	238	2 Q8WTC6	Q8Wtc6 aequorea ma
12	1065	83.9	238	2 Q8WTC4	Q8Wtc4 aequorea ma
13	1063	83.7	238	2 Q8WTD0	Q8Wtd0 aequorea ma
14	1062	83.6	238	2 Q8WTC8	Q8Wtc8 aequorea ma
15	1062	83.6	238	2 Q8WTC9	Q8Wtc9 aequorea ma
16	1060	83.5	238	2 Q8WTC7	Q8Wtc7 aequorea ma
17	1058	83.3	238	2 Q8WTC5	Q8Wtc5 aequorea ma
18	602	47.4	234	2 Q6RYS7	Q6Rys7 phialidium
19	463	36.5	225	2 Q6RYS5	Q6Rys5 anthomedusa
20	429	33.8	262	2 Q6RYS6	Q6Rys6 anthomedusa
21	255.5	20.1	225	2 Q95UA7	Q95ua7 montastraea
22	255.5	20.1	225	2 Q7Z0W5	Q7Z0w5 montastraea
23	250.5	19.7	225	2 Q7Z0W9	Q7Z0w9 montastraea
24	245	19.3	225	2 Q963F5	Q963f5 montastraea
25	244.5	19.3	225	2 Q6USK3	Q6usk3 montastraea
26	243.5	19.2	226	2 Q8T6U0	Q8t6u0 dendronept
27	241.5	19.0	266	2 Q9U6V3	Q9u6v3 clavularia
28	241	19.0	225	2 Q8I6J8	Q8i6j8 trachyphyl
29	241	19.0	230	2 Q66PW1	Q66pw1 scolymia cu
30	239.5	18.9	225	2 Q8T5F1	Q8t5f1 montastraea
31	238	18.7	232	2 Q6RYS4	Q6Rys4 anthomedusa

32	233	18.3	225	2	Q66ND3	Q66nd3 montastraea
33	232.5	18.3	224	2	Q8MU48	Q8mu48 montastraea
34	231	18.2	225	2	Q7Z0W4	Q7z0w4 montastraea
35	224	17.6	231	2	Q66PV5	Q66pv5 acropora mi
36	223.5	17.6	231	2	Q6R8F5	Q6r8f5 astrangia 1
37	221.5	17.4	221	2	Q66PV1	Q66pv1 acropora te
38	218	17.2	231	2	Q66PV3	Q66pv3 acropora no
39	216	17.0	219	2	Q6R8F4	Q6r8f4 astrangia 1
40	216	17.0	232	2	Q6I7B1	Q6i7b1 acropora sp
41	215	17.0	259	2	Q8MMA2	Q8mma2 agaricia fr
42	215.5	17.0	227	2	Q66ND5	Q66nd5 montastraea
43	214.5	16.9	218	2	Q6I7B2	Q6i7b2 fungia conc
44	213	16.8	227	2	Q66ND4	Q66nd4 montastraea
45	213	16.8	231	2	Q66PV4	Q66pv4 acropora no

ALIGNMENTS

RESULT 1
Q8GHE2 PRELIMINARY; PRT; 238 AA.
AC Q8GHE2
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescence protein.
GN Name=2289Gfp;
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM2289;
RA Koranyi P., Berenyi M., Burg K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF324408; AAN86140.1; -
DR HSSP; P42212; IGFL.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR011584; GFP_related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPUCESCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 26914 MW; F84840F1F9064018 CRC64;

Query Match	98.1%;	Score	1246;	DB	2;	Length	238;
Best Local Similarity	97.9%;	Pred. No.	6.3e-93;				
Matches	233;	Conservative	2;	Mismatches	3;	Indels	0;
Gaps	0;						
Qy	1	MSGEELFTAVPILVELDGDVNGHKFSVSGEGEDVTYGYKLTILKFKICTTGKLPVHPPTL	60				
Db	1	MSGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGYKLTILKFKICTTGKLPVHPPTL	60				
Qy	61	VTFSTVGVCFSYPPDHMKRDFPKSAMPGGYVQORTIFFKDDGNYKTRAEVKEGDTLV	120				
Db	61	VTFSTVGVCFSYPPDHMKRDFPKSAMPGGYVQERTIFFKDDGNYKTRAEVKEGDTLV	120				
Qy	121	NRIELKGIDPKEDGNILGHKLEYNYNHSHNYYIMADKOKNGIKVNFKIRHNIEDGSVQLAD	180				
Db	121	NRIELKGIDPKEDGNILGHKLEYNYNHSHNYYIMADKOKNGIKVNFKIRHNIEDGSVQLAD	180				
Qy	181	YYQNTPIIDGPPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDLYK	238				
Db	181	HYQNTPIIDGPPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDLYK	238				
RESULT 2							
GFP_AEQVI							
ID -GFP_AEQVI							
AC	P42212;	Q17104;	Q27903;				

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Green fluorescent protein.
GN Names:GFP;
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP MEDLINE=92175527; PubMed=1347277; DOI=10.1016/0378-1119(92)90691-H;
RX Prasher D.C., Sckenrode V.K., Ward W.W., Prendergast F.G.,
RA Cormier M.J.;
RT "Primary structure of the Aequorea victoria green-fluorescent
RT protein.";
RL Gene 111:229-233(1992).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=94185810; PubMed=8137953; DOI=10.1016/0014-5793(94)80472-9;
RX Inouye S., Tsuji F.I.;
RT "Aequorea green fluorescent protein. Expression of the gene and
RT fluorescence characteristics of the recombinant protein.";
RL FEBS Lett. 341:277-280(1994).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=97299832; PubMed=9154981; DOI=10.1023/A:1005740823703;
RX Rouwendal G.J.A., Mendes O., Wolbert E.J.H., de Boer A.D.;
RT "Enhanced expression in tobacco of the gene encoding green fluorescent
RT protein by modification of its codon usage.";
RL Plant Mol. Biol. 33:989-999(1997).
[4]
RN CHROMOPHORE.
RP MEDLINE=93192221; PubMed=8448132;
RX Cody C.W., Prasher D.C., Westler W.M., Prendergast F.G., Ward W.W.;
RA "Chemical structure of the hexapeptide chromophore of the Aequorea
RT green-fluorescent protein.";
RL Biochemistry 32:1212-1218(1993).
[5]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RP MEDLINE=96355665; PubMed=8703075;
RX Ormoe M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y.,
RA Remington S.J.;
RT "Crystal structure of the Aequorea victoria green fluorescent
RT protein.";
RL Science 273:1392-1395(1996).
[6]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RP MEDLINE=98294543; PubMed=9631087;
RX Yang F., Moss L.G., Phillips G.N. Jr.;
RT "The molecular structure of green fluorescent protein.";
RL Nat. Biotechnol. 14:1246-1251(1996).
[7]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF MUTANT WITH YELLOW EMISSION.
RP MEDLINE=98455509; PubMed=9782051; DOI=10.1016/S0969-2126(98)00127-0;
RX Wachter R.M., Elsiger M.A., Kallio K., Hanson G.T., Remington S.J.;
RA "Structural basis of spectral shifts in the yellow-emission variants
RT of green fluorescent protein.";
RL Structure 6:1267-1277(1998).
[8]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RP MEDLINE=99238303; PubMed=10220315; DOI=10.1021/bi9902182;
RX Elsiger M.A., Wachter R.M., Hanson G.T., Kallio K., Remington S.J.;
RA "Structural and spectral response of green fluorescent protein
RT variants to changes in pH.";
RL Biochemistry 38:5296-5301(1999).
CC -I- FUNCTION: Energy-transfer acceptor. Its role is to transduce the
CC blue chemiluminescence of the protein aequorin into green
CC fluorescent light by energy transfer. Fluoresces in vivo upon
CC receiving energy from the Ca(2+)-activated photoprotein aequorin.
CC Absorbs light maximally at 395 nm and exhibits a smaller
CC absorbance peak at 470 nm. The fluorescence emission spectrum
CC peaks at 509 nm with a shoulder at 540 nm.

CC -I- SUBUNIT: Monomer.
CC -I- TISSUE SPECIFICITY: Photocytes.
CC -I- PTM: Contains a covalently attached chromophore, which is composed
CC of modified amino acid residues. The chromophore is formed upon
CC cyclization of the residues Ser-dehydrotyr-Gly.
CC -I- BIOTECHNOLOGY: Has become a useful and ubiquitous tool for making
CC chimeric proteins of GFP linked to other proteins where it
CC functions as a fluorescent protein tag. GFP tolerates N-and C-
CC terminal fusion to a broad variety of proteins. It has been
CC expressed in bacteria, yeast, slime mold, plants, Drosophila,
CC zebrafish, and in mammalian cells. As a noninvasive fluorescent
CC marker in living cells, it allows for a wide range of applications
CC where it may function as a cell lineage tracer, reporter of gene
CC expression, or as a measure of protein-protein interactions.
CC -I- DATABASE: NAME=Protein Spotlight; NOTE=Issue 11 of June 2001;
CC WWW="http://www.expasy.org/spotlight/articles/sptlt011.html".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M62654; AAA27722.1; -;
DR EMBL; M62653; AAA27721.1; -;
DR EMBL; L29345; AAA58246.1; -;
DR EMBL; X96418; CAA65278.1; -;
DR PIR; JS0692; JQ1514.
DR PDB; 1B9C; X-ray; A/B/C/D=1-238.
DR PDB; 1BFP; X-ray; @=1-238.
DR PDB; 1C4F; X-ray; A=1-238.
DR PDB; 1CV7; X-ray; A=1-228.
DR PDB; 1EMA; X-ray; @=-.
DR PDB; 1EMB; X-ray; @=1-238.
DR PDB; 1ENC; X-ray; A/B/C/D=1-238.
DR PDB; 1EME; X-ray; @=1-238.
DR PDB; 1EMF; X-ray; @=1-238.
DR PDB; 1ENG; X-ray; A=1-238.
DR PDB; 1ENL; X-ray; @=1-238.
DR PDB; 1ENM; X-ray; @=1-238.
DR PDB; 1F09; X-ray; A=1-238.
DR PDB; 1F0B; X-ray; A=1-238.
DR PDB; 1GFL; X-ray; A/B=1-238.
DR PDB; 1H6R; X-ray; A/B/C=1-238.
DR PDB; 1HGT; X-ray; A/B/C/D=1-238.
DR PDB; 1HUY; X-ray; A=1-238.
DR PDB; 1JBY; X-ray; A=1-238.
DR PDB; 1JBZ; X-ray; A=1-238.
DR PDB; 1JC0; X-ray; A/B/C=1-238.
DR PDB; 1JC1; X-ray; A/B/C=1-237.
DR PDB; 1KPS; X-ray; A=1-238.
DR PDB; 1KYP; X-ray; A=1-238.
DR PDB; 1KYR; X-ray; A=1-238.
DR PDB; 1MYW; X-ray; A=1-238.
DR PDB; 1Q4A; X-ray; A=1-238.
DR PDB; 1Q4B; X-ray; A=1-238.
DR PDB; 1Q4C; X-ray; A=1-238.
DR PDB; 1Q4D; X-ray; A=1-238.
DR PDB; 1Q4E; X-ray; A=1-238.
DR PDB; 1Q73; X-ray; A=1-238.
DR PDB; 1QXT; X-ray; A=1-229.
DR PDB; 1QX3; X-ray; A=1-229.
DR PDB; 1QXF; X-ray; A=1-229.
DR PDB; 1QXO; X-ray; A=1-238.
DR PDB; 1QXQ; X-ray; A=1-238.
DR PDB; 1YFP; X-ray; @=1-238.
DR PDB; 2EMN; X-ray; @=1-238.
DR PDB; 2EMO; X-ray; @=1-238.

```
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP; 1.
DR ProDom; PD013756; Green_fl_protein; 1.
KW 3D-structure; Direct protein sequencing; Luminescence.
FT CROSSLINK 65 67
FT MOD_RES 66 66 5-imidazolinone (Ser-Gly).
FT VAR 66 66 (Z)-2,3-didehydrotyrosine.
FT VARIANT 100 100 F -> Y.
FT VARIANT 108 108 T -> S.
FT VARIANT 141 141 L -> M.
FT VARIANT 219 219 V -> I.
FT CONFLICT 2 2 S -> G (in Ref. 3).
FT CONFLICT 25 25 H -> Q (in Ref. 2).
FT CONFLICT 80 80 Q -> R (in Ref. 3).
FT CONFLICT 157 157 Q -> P (in Ref. 2).
FT CONFLICT 172 172 E -> K (in Ref. 2).
FT TURN 3 3
FT HELIX 4 8
FT STRAND 11 22
FT TURN 23 24
FT STRAND 25 26
FT STRAND 37 39
FT HELIX 40 40
FT TURN 41 48
FT STRAND 49 50
FT TURN 57 60
FT HELIX 61 63
FT TURN 69 71
FT STRAND 73 73
FT TURN 76 81
FT HELIX 83 86
FT TURN 90 90
FT STRAND 92 100
FT TURN 101 102
FT STRAND 105 115
FT TURN 116 117
FT STRAND 118 128
FT TURN 132 133
FT TURN 135 139
FT STRAND 141 141
FT STRAND 148 155
FT HELIX 156 158
FT TURN 159 159
FT STRAND 160 171
FT TURN 172 173
FT STRAND 176 187
FT TURN 199 208
FT TURN 211 212
FT STRAND 217 227
SQ SEQUENCE 238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;

Query Match 97.8%; Score 1242; DB 1; Length 238;
Best Local Similarity 97.5%; Pred. No. 1.3e-92;
Matches 232; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPLVELDGVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPLVELDGVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVPWPTL 60
Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNHNVYIMADKQNGIKVNFKIRHNIEDGSGVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNHNVYIMADKQNGIKVNFKIRHNIEDGSGVQLAD 180
Qy 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 YQOQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 4
Q8GHE4 PRELIMINARY; PRT; 238 AA.
AC Q8GHE4
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescence protein.
GN Names=375Gfp;
OS Azomonas agilis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azomonas.
OX NCBI_TaxID=116849;
[1]
RN SEQUENCE FROM N.A.
RA Koranyi P., Berenyi M., Burg K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AF324407; AAN86137.1; -.
DR HSP; P42212; 1B9C.
GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR011584; GFP_related.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP; 1.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;

Query Match 97.8%; Score 1242; DB 2; Length 238;
Best Local Similarity 97.5%; Pred. No. 1.3e-92;
Matches 232; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPLVELDGVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPLVELDGVNGHKFSVSGEGSDVYTKLTLKFKICTTGKLPVPWPTL 60
Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNYNHNVYIMADKQNGIKVNFKIRHNIEDGSGVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNYNHNVYIMADKQNGIKVNFKIRHNIEDGSGVQLAD 180
Qy 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 YQOQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 3
Q71RY9 PRELIMINARY; PRT; 238 AA.
AC Q71RY9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Green fluorescence protein.
GN Name=289Gfp;
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
[1]
RN SEQUENCE FROM N.A.
RA Koranyi P., Berenyi M., Burg K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AF324407; AAN86139.1; -.
DR HSP; P42212; 1B9C.
GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR011584; GFP_related.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP; 1.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;
```

```
DR Pfam: PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR PRODOM; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 26902 MW; 15F59B9C5B4F6B89 CRC64;

Query Match          97.7%; Score 1241; DB 2; Length 238;
Best Local Similarity 97.5%; Pred. No. 1.9e-92;
Matches 232; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGEDVTVYGLTKLFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPILVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60

QY 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180

QY 181 YQOQNTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238
DB 181 YQOQNTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 5
Q8GHE3 PRELIMINARY; PRT; 238 AA.
AC Q8GHE3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescence protein.
GN Name=85Gfp;
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RA Koranyi P., Berenyi M., Burg K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE324406; AAN86138.1; -
DR HSSP; P42212; IGFL.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR011584; GFP related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR PRODOM; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 26887 MW; B0E1616BD2AF6188 CRC64;

Query Match          97.6%; Score 1240; DB 2; Length 238;
Best Local Similarity 97.5%; Pred. No. 1.9e-92;
Matches 232; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGEDVTVYGLTKLFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPILVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60

QY 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180

QY 181 YQOQNTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238
DB 181 YQOQNTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238
```

```
RESULT 6
Q93125 PRELIMINARY; PRT; 238 AA.
ID Q93125;
AC Q93125;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein mutant 3.
GN Name=GFP;
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97195776; PubMed=8707053; DOI=10.1016/0378-1119(95)00685-0;
RA Cormack B.P., Valdivia R.H., Falkow S.;
RT "FACS-optimized mutants of the green fluorescent protein (GFP).";
RL Gene 173:33-38 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97195776; PubMed=9043107;
RA Cormack B.P., Bertram G., Egerton M., Gow N.A.R., Falkow S.,
RA Brown A.J.P.;
RT "Yeast-enhanced green fluorescent protein (yEGFP) a reporter of gene
expression in Candida albicans.";
RL Microbiology 143:303-311 (1997).
DR EMBL; U73901; AAB18957.1; -
DR PDB; 2YFP; X-ray; A=1-238.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR011584; GFP related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
SQ SEQUENCE 238 AA; 26840 MW; A28622809A9DEA60 CRC64;

Query Match          97.2%; Score 1235; DB 2; Length 238;
Best Local Similarity 96.6%; Pred. No. 4.9e-92;
Matches 230; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEGEDVTVYGLTKLFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPILVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60

QY 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNHNHVIYIMADKQKNGIKVNFKIRHNIEDGSVOLAD 180

QY 181 YQOQNTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238
DB 181 YQOQNTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 7
Q17105 PRELIMINARY; PRT; 238 AA.
ID Q17105;
AC Q17105;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein (Fragment).
GN Name=GFP;
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
```



```
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Green fluorescent protein.
GN Name=GFPxm;
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qin Y.X., Pang S.Q.,
RA Li S.J., Xia N.S.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY013824; AA02062.1; -
DR EMBL; AY013824; AA02062.1; -
DR HSSP; P42212; 1BFP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
SQ SEQUENCE 238 AA; 27049 MW; 8185D0E5E529012B CRC64;

Query Match 85.5%; Score 1079; DB 2; Length 238;
Best Local Similarity 81.5%; Pred. No. 2.1e-79;
Matches 194; Conservative 21; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVPILVELDGVNKHKFSVSGEGDVTYKLTKEICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGIVPVLIELDGVNKHKFSVSGEGDADYKLEIKFICTTGKLPVWPPTL 60

QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKTRAEVKFEGDTLV 120
DB 61 VTTFSGVQICFARYPEHMKNDFFKSAMPEGYQERTIFFQDDGKYKTRGEVKFEGDTLV 120

QY 121 NRIELKGIDPKEDGNILGHKLEYNHSHNVIYIMADKQKGIKVNFKIRHNIEGGSVOLAD 180
DB 121 NRIELKGMPDKEDGNILGHKLEYNHSHNVIYIMPDKANGLKVNFKIRHNIEGGSVOLAD 180

QY 181 YQONTPIIDGPVLLPDNHYLSTQSALSQDPNEKRDHMLLEFVTAAGTHGMDELYK 238
DB 181 HYQTNVPLGDPVLLPINHYLSTQTAISKDRNETRDHMLVLEFVTAAGTHGMDELYK 238

RESULT 11
Q8WTC6 PRELIMINARY; PRT; 238 AA.
ID Q8WTC6;
AC Q8WTC6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN Name=GFP;
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435433; AAL33918.1; -
DR HSSP; P42212; 1KYP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
SQ SEQUENCE 238 AA; 26956 MW; 75521EAF0CEBA73A CRC64;

Query Match 83.9%; Score 1065; DB 2; Length 238;
Best Local Similarity 81.5%; Pred. No. 2.9e-78;
Matches 194; Conservative 18; Mismatches 26; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVPILVELDGVNKHKFSVSGEGDVTYKLTKEICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGIVPVLIELDGVNKHKFSVSGEGDADYKLEIKFICTTGKLPVWPPTL 60

QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKTRAEVKFEGDTLV 120
DB 61 VTTFSGVQICFARYPEHMKNDFFKSAMPEGYQERTIFFQDDGKYKTRGEVKFEGDTLV 120

QY 121 NRIELKGIDPKEDGNILGHKLEYNHSHNVIYIMADKQKGIKVNFKIRHNIEGGSVOLAD 180
DB 121 NRIELKGMPDKEDGNILGHKLEYNHSHNVIYIMPDKANGLKVNFKIRHNIEGGSVOLAD 180

QY 181 YQONTPIIDGPVLLPDNHYLSTQSALSQDPNEKRDHMLLEFVTAAGTHGMDELYK 238
DB 181 HYQTNVPLGDPVLLPINHYLSTQTAISKDRNETRDHMLVLEFVTAAGTHGMDELYK 238

RESULT 13
Q8WTD0 PRELIMINARY; PRT; 238 AA.
ID Q8WTD0;
AC Q8WTD0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN Name=GFP;
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435431; AAL33916.1; -
DR HSSP; P42212; 1KYP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
SQ SEQUENCE 238 AA; 27015 MW; 6B8FD75B88926903 CRC64;

Query Match 84.5%; Score 1073; DB 2; Length 238;
Best Local Similarity 81.3%; Pred. No. 6.4e-79;
Matches 193; Conservative 21; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVPILVELDGVNKHKFSVSGEGDVTYKLTKEICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGIVPVLIELDGVNKHKFSVSGEGDADYKLEIKFICTTGKLPVWPPTL 60
```

```

[1]
RN  SEQUENCE FROM N.A.
RP  Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA  Li S.J., Xia N.S.;
RL  Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF435427; AAL33912.1; -
DR  HSP; P42212; IKYP.
DR  GO; GO:0006091; P:energy pathways; IEA.
DR  Pfam; PF01353; GFP; 1.
DR  PRINTS; PR01229; GFP; 1.
DR  ProDom; PD013756; Green fl protein; 1.
SQ  SEQUENCE 238 AA; 26997 MW; 5F80A192173CB84D CRC64;

Query Match      83.7%; Score 1063; DB 2; Length 238;
Best Local Similarity 80.3%; Pred. No. 4.2e-78;
Matches 191; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY  1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEQDVYTGKLTIKFTCTTGKLPVPWPTL 60
DB  1 MSKGEELFTGIVPVLIELDGVNGHKFSVRGEGDADYGKLEIKFTCTTGKLPVPWPTL 60
QY  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB  61 VTTLGYGIQCFARYPEHMKNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120
QY  121 NRLEKGMDFKEDGNILGHKLEYNYNHSHVYIMADKQNGIKYNFKIRHNIEDGVSQVLAD 180
DB  121 NRLEKGMDFKEDGNILGHKLEYNFNHSHVYIMADKANNGLKYNFKIRHNIEGGVQVLAD 180
QY  181 YYQONTPILOGPVLPPDNHLYSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
DB  181 HYQTNVPLGDPVLIPINHLYSLQTAISKORNETRDMHNVLEFFSACGHTHGMDELYK 238

[1]
RN  SEQUENCE FROM N.A.
RP  Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA  Li S.J., Xia N.S.;
RL  Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF435428; AAL33913.1; -
DR  HSP; P42212; IKYP.
DR  GO; GO:0006091; P:energy pathways; IEA.
DR  Pfam; PF01353; GFP; 1.
DR  PRINTS; PR01229; GFP; 1.
DR  ProDom; PD013756; Green fl protein; 1.
SQ  SEQUENCE 238 AA; 27031 MW; 5F80A19C19DC584D CRC64;

Query Match      83.6%; Score 1062; DB 2; Length 238;
Best Local Similarity 80.3%; Pred. No. 5e-78;
Matches 191; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY  1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEQDVYTGKLTIKFTCTTGKLPVPWPTL 60
DB  1 MSKGEELFTGIVPVLIELDGVNGHKFSVRGEGDADYGKLEIKFTCTTGKLPVPWPTL 60
QY  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB  61 VTTLGYGIQCFARYPEHMKNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120
QY  121 NRLEKGMDFKEDGNILGHKLEYNYNHSHVYIMADKQNGIKYNFKIRHNIEDGVSQVLAD 180
DB  121 NRLEKGMDFKEDGNILGHKLEYNFNHSHVYIMADKANNGLKYNFKIRHNIEGGVQVLAD 180
QY  181 YYQONTPILOGPVLPPDNHLYSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
DB  181 HYQTNVPLGDPVLIPINHLYSLQTAISKORNETRDMHNVLEFFSACGHTHGMDELYK 238

[1]
RN  SEQUENCE FROM N.A.
RP  Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA  Li S.J., Xia N.S.;
RL  Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF435429; AAL33914.1; -
DR  HSP; P42212; IKYP.
DR  GO; GO:0006091; P:energy pathways; IEA.
DR  Pfam; PF01353; GFP; 1.
DR  PRINTS; PR01229; GFP; 1.
DR  ProDom; PD013756; Green fl protein; 1.
SQ  SEQUENCE 238 AA; 27047 MW; 5F80A18FA1E7C84D CRC64;

Query Match      83.6%; Score 1062; DB 2; Length 238;
Best Local Similarity 80.3%; Pred. No. 5e-78;
Matches 191; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY  1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEQDVYTGKLTIKFTCTTGKLPVPWPTL 60
DB  1 MSKGEELFTGIVPVLIELDGVNGHKFSVRGEGDADYGKLEIKFTCTTGKLPVPWPTL 60
QY  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB  61 VTTLGYGIQCFARYPEHMKNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120
QY  121 NRLEKGMDFKEDGNILGHKLEYNYNHSHVYIMADKQNGIKYNFKIRHNIEDGVSQVLAD 180
DB  121 NRLEKGMDFKEDGNILGHKLEYNFNHSHVYIMADKANNGLKYNFKIRHNIEGGVQVLAD 180
QY  181 YYQONTPILOGPVLPPDNHLYSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
DB  181 HYQTNVPLGDPVLIPINHLYSLQTAISKORNETRDMHNVLEFFSACGHTHGMDELYK 238
```

```

[1]
RN  SEQUENCE FROM N.A.
RP  Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA  Li S.J., Xia N.S.;
RL  Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF435427; AAL33912.1; -
DR  HSP; P42212; IKYP.
DR  GO; GO:0006091; P:energy pathways; IEA.
DR  Pfam; PF01353; GFP; 1.
DR  PRINTS; PR01229; GFP; 1.
DR  ProDom; PD013756; Green fl protein; 1.
SQ  SEQUENCE 238 AA; 26997 MW; 5F80A192173CB84D CRC64;

Query Match      83.7%; Score 1063; DB 2; Length 238;
Best Local Similarity 80.3%; Pred. No. 4.2e-78;
Matches 191; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY  1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEQDVYTGKLTIKFTCTTGKLPVPWPTL 60
DB  1 MSKGEELFTGIVPVLIELDGVNGHKFSVRGEGDADYGKLEIKFTCTTGKLPVPWPTL 60
QY  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB  61 VTTLGYGIQCFARYPEHMKNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120
QY  121 NRLEKGMDFKEDGNILGHKLEYNYNHSHVYIMADKQNGIKYNFKIRHNIEDGVSQVLAD 180
DB  121 NRLEKGMDFKEDGNILGHKLEYNFNHSHVYIMADKANNGLKYNFKIRHNIEGGVQVLAD 180
QY  181 YYQONTPILOGPVLPPDNHLYSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
DB  181 HYQTNVPLGDPVLIPINHLYSLQTAISKORNETRDMHNVLEFFSACGHTHGMDELYK 238

[1]
RN  SEQUENCE FROM N.A.
RP  Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA  Li S.J., Xia N.S.;
RL  Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF435428; AAL33913.1; -
DR  HSP; P42212; IKYP.
DR  GO; GO:0006091; P:energy pathways; IEA.
DR  Pfam; PF01353; GFP; 1.
DR  PRINTS; PR01229; GFP; 1.
DR  ProDom; PD013756; Green fl protein; 1.
SQ  SEQUENCE 238 AA; 27031 MW; 5F80A19C19DC584D CRC64;

Query Match      83.6%; Score 1062; DB 2; Length 238;
Best Local Similarity 80.3%; Pred. No. 5e-78;
Matches 191; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY  1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEQDVYTGKLTIKFTCTTGKLPVPWPTL 60
DB  1 MSKGEELFTGIVPVLIELDGVNGHKFSVRGEGDADYGKLEIKFTCTTGKLPVPWPTL 60
QY  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB  61 VTTLGYGIQCFARYPEHMKNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120
QY  121 NRLEKGMDFKEDGNILGHKLEYNYNHSHVYIMADKQNGIKYNFKIRHNIEDGVSQVLAD 180
DB  121 NRLEKGMDFKEDGNILGHKLEYNFNHSHVYIMADKANNGLKYNFKIRHNIEGGVQVLAD 180
QY  181 YYQONTPILOGPVLPPDNHLYSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
DB  181 HYQTNVPLGDPVLIPINHLYSLQTAISKORNETRDMHNVLEFFSACGHTHGMDELYK 238

[1]
RN  SEQUENCE FROM N.A.
RP  Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA  Li S.J., Xia N.S.;
RL  Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF435429; AAL33914.1; -
DR  HSP; P42212; IKYP.
DR  GO; GO:0006091; P:energy pathways; IEA.
DR  Pfam; PF01353; GFP; 1.
DR  PRINTS; PR01229; GFP; 1.
DR  ProDom; PD013756; Green fl protein; 1.
SQ  SEQUENCE 238 AA; 27047 MW; 5F80A18FA1E7C84D CRC64;

Query Match      83.6%; Score 1062; DB 2; Length 238;
Best Local Similarity 80.3%; Pred. No. 5e-78;
Matches 191; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

QY  1 MSKGEELFTAVVPILVELDGVNGHKFSVSGEQDVYTGKLTIKFTCTTGKLPVPWPTL 60
DB  1 MSKGEELFTGIVPVLIELDGVNGHKFSVRGEGDADYGKLEIKFTCTTGKLPVPWPTL 60
QY  61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB  61 VTTLGYGIQCFARYPEHMKNDFFKSAMPEGYIQERTIFFQDDGKYKTRGEVKFEGDTLV 120
QY  121 NRLEKGMDFKEDGNILGHKLEYNYNHSHVYIMADKQNGIKYNFKIRHNIEDGVSQVLAD 180
DB  121 NRLEKGMDFKEDGNILGHKLEYNFNHSHVYIMADKANNGLKYNFKIRHNIEGGVQVLAD 180
QY  181 YYQONTPILOGPVLPPDNHLYSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
DB  181 HYQTNVPLGDPVLIPINHLYSLQTAISKORNETRDMHNVLEFFSACGHTHGMDELYK 238
```

Search completed: April 25, 2005, 15:46:41
Job time : 182 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 15:21:40 ; Search time 167 Seconds
(without alignments)

551.191 Million cell updates/sec

Title: US-10-620-099-2

Perfect score: 1270

Sequence: 1 MSKGEELFTAVVPILVELDG.....VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1270	100.0	238	2	AAW52313 Aequorea
2	1267	99.8	238	2	AAW52328 Aequorea
3	1267	99.8	238	2	AAW52330 Aequorea
4	1266	99.7	238	2	AAW52333 Aequorea
5	1266	99.7	238	2	AAW52315 Aequorea
6	1265	99.6	238	2	AAW52314 Aequorea
7	1265	99.6	238	2	AAW52336 Aequorea
8	1265	99.6	238	2	AAW52351 Aequorea
9	1265	99.6	238	2	AAW52334 Aequorea
10	1265	99.6	238	2	AAW52335 Aequorea
11	1265	99.6	238	2	AAW52331 Aequorea
12	1265	99.6	238	2	AAW52337 Aequorea
13	1265	99.6	238	2	AAW52349 Aequorea
14	1265	99.6	238	2	AAW52346 Aequorea
15	1265	99.6	238	2	AAW52324 Aequorea
16	1265	99.6	238	2	AAW52338 Aequorea
17	1265	99.6	238	2	AAW52344 Aequorea
18	1264	99.5	238	2	AAW52332 Aequorea
19	1264	99.5	238	2	AAW52347 Aequorea
20	1264	99.5	238	2	AAW52339 Aequorea
21	1264	99.5	238	2	AAW52350 Aequorea
22	1264	99.5	238	2	AAW52340 Aequorea
23	1263	99.4	238	2	AAW52348 Aequorea
24	1263	99.4	238	2	AAW52345 Aequorea
25	1263	99.4	238	2	AAW52341 Aequorea

26	1262	99.4	238	2	AAW52342 Aequorea
27	1262	99.4	238	2	AAW52323 Aequorea
28	1261	99.3	238	2	AAW52343 Aequorea
29	1261	99.3	238	2	AAW52325 Aequorea
30	1260	99.2	238	2	AAW52317 Aequorea
31	1260	99.2	238	2	AAW52316 Aequorea
32	1257	99.0	238	2	AAW52329 Aequorea
33	1256	98.9	238	2	AAW52322 Aequorea
34	1256	98.9	238	2	AAW52321 Aequorea
35	1253	98.7	238	2	AAW52319 Aequorea
36	1250	98.4	238	2	AAW52318 Aequorea
37	1249	98.3	238	2	AAW52320 Aequorea
38	1246	98.1	238	2	AAW05304 Aequorea
39	1246	98.1	238	2	AAW24232 Aequorea
40	1246	98.1	238	2	AAW76105 Aequorea
41	1246	98.1	238	2	AAW40479 Aequorea
42	1246	98.1	238	2	AAW76371 Aequorea
43	1246	98.1	238	4	AAW73552 Aequorea
44	1246	98.1	238	5	AAE16038 Aequorea
45	1246	98.1	238	5	ABG32365 Aequorea

ALIGNMENTS

RESULT 1
AAW52313
ID AAW52313 standard; protein; 238 AA.
XX
AC AAW52313;
XX
DT 10-JUL-1998 (first entry)
XX
DE Aequorea green fluorescent protein.
XX
KW Green fluorescent protein; engineered fluorescent protein;
KW reporter molecule; immunological assay; protein-protein interaction;
KW fluorescence resonance energy transfer system; FRET system.
XX
OS Aequorea victoria.
XX
FN WO9806737-A1.
XX
PD 19-FEB-1998.
XX
PF 15-AUG-1997; 97WO-US014593.
XX
PR 16-AUG-1996; 96US-0024050P.
PR 30-AUG-1996; 96US-00706408.
XX
PA (REGC) UNIV CALIFORNIA.
PA (UYOR-) UNIV OREGON.
XX
PA (AURO-) AURORE BIOSCIENCES.
XX
PI Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
XX
DR WPI; 1998-159454/14.
XX
N-PSDB; AAV19946.
XX
PT Nucleic acid encoding mutant green fluorescent proteins having longer
PT wavelength emission - used as markers for probes and as components of
PT fluorescent resonant energy transfer systems, also related vectors and
PT transformants.
XX
PS Claim 1; Page 53; 120pp; English.
XX
CC This sequence is the wild type Aequorea victoria green fluorescent
CC protein. The nucleic acid was mutated to produce the nucleic acids (I) of
CC the invention, which encode functional, engineered fluorescent proteins
CC (II) having largely the same sequence as Aequorea green fluorescent
CC protein (GFP) but differing by at least the substitution T203X (X = H, Y,
CC W or F) and having fluorescent properties different from GFP. (II) are
CC useful as reporter molecules in immunological or hybridisation assays,

CC for monitoring proteins in cells and detecting induction of
 CC transcription. They are also useful as components of FRET (fluorescence
 CC resonance energy transfer) systems, e.g. for detecting protein-protein
 CC interactions, cleavage of substrates and changes in potential across a
 CC membrane, and further for making fluorescent substrates for protein
 CC kinase. (II) are easily distinguished from known green and blue
 CC fluorescent proteins, so allow simultaneous measurements of at least 2
 CC proteins in a cell. They emit, and are excited, at relatively long
 CC wavelengths where phototoxicity and auto-fluorescence are reduced
 XX Sequence 238 AA;

Query Match 100.0%; Score 1270; DB 2; Length 238;
 Best Local Similarity 100.0%; Pred. No. 6.5e-123;
 Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKGEELFTAVVPILVELDGVNGHKFVSVEGEGDVTYGLKLFKICTTGKLPVPMPTL 60
 DB 1 MSKGEELFTAVVPILVELDGVNGHKFVSVEGEGDVTYGLKLFKICTTGKLPVPMPTL 60
 QY 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
 DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
 QY 121 NRIELKGIDPKEDGNILGHKLEYNYNHSHVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
 DB 121 NRIELKGIDPKEDGNILGHKLEYNYNHSHVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
 QY 181 YQQNTPILDGPVLLPDNHYLSQTSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
 DB 181 YQQNTPILDGPVLLPDNHYLSQTSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 2
 AAW52328
 ID AAW52328 standard; protein; 238 AA.
 XX AAW52328;
 DT 10-JUL-1998 (first entry)
 XX Engineered green fluorescent protein S65T.
 XX Green fluorescent protein; engineered fluorescent protein; mutein;
 KW reporter molecule; immunological assay; protein-protein interaction;
 KW fluorescence resonance energy transfer system; FRET system.
 XX Synthetic.
 OS Aequorea victoria.
 XX Key Location/Qualifiers
 FT Misc-difference 65
 FT /label= S65T
 FT /note= "Ser to Thr mutation"
 XX WO9806737-A1.
 XX 19-FEB-1998.
 XX 15-AUG-1997; 97WO-US014593.
 XX 16-AUG-1996; 96US-0024050P.
 XX 30-AUG-1996; 96US-00706408.
 XX (REGC) UNIV CALIFORNIA.
 PA (UWOR-) UNIV OREGON
 PA (AURO-) AURORE BIOSCIENCES.
 XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
 XX WPI; 1998-159454/14.
 XX Nucleic acid encoding mutant green fluorescent proteins having longer

PT wavelength emission - used as markers for probes and as components of
 PT fluorescent resonant energy transfer systems, also related vectors and
 PT transformants.

Claim 5; Page; 120pp; English.

CC This sequence is an engineered Aequorea victoria green fluorescent
 CC protein (GFP) of the invention, created from the sequence shown in
 CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
 CC invention, which encode functional, engineered fluorescent proteins (II)
 CC having largely the same sequence as Aequorea green fluorescent protein
 CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
 CC and having fluorescent properties different from GFP. (II) are useful as
 CC reporter molecules in immunological or hybridisation assays, for
 CC monitoring proteins in cells and detecting induction of transcription.
 CC They are also useful as components of FRET (fluorescence resonance energy
 CC transfer) systems, e.g. for detecting protein-protein interactions,
 CC cleavage of substrates and changes in potential across a membrane, and
 CC further for making fluorescent substrates for protein kinase. (II) are
 CC easily distinguished from known green and blue fluorescent proteins, so
 CC allow simultaneous measurements of at least 2 proteins in a cell. They
 CC emit, and are excited, at relatively long wavelengths where phototoxicity
 CC and auto-fluorescence are reduced
 XX Sequence 238 AA;

Query Match 99.8%; Score 1267; DB 2; Length 238;
 Best Local Similarity 99.6%; Pred. No. 1.3e-122;
 Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKGEELFTAVVPILVELDGVNGHKFVSVEGEGDVTYGLKLFKICTTGKLPVPMPTL 60
 DB 1 MSKGEELFTAVVPILVELDGVNGHKFVSVEGEGDVTYGLKLFKICTTGKLPVPMPTL 60
 QY 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
 DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNKYKTRAEVKFEGDTLV 120
 QY 121 NRIELKGIDPKEDGNILGHKLEYNYNHSHVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
 DB 121 NRIELKGIDPKEDGNILGHKLEYNYNHSHVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
 QY 181 YQQNTPILDGPVLLPDNHYLSQTSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238
 DB 181 YQQNTPILDGPVLLPDNHYLSQTSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 3
 AAW52330
 ID AAW52330 standard; protein; 238 AA.
 XX AAW52330;
 XX 10-JUL-1998 (first entry)
 XX Engineered green fluorescent protein S65A.
 XX Green fluorescent protein; engineered fluorescent protein; mutein;
 KW reporter molecule; immunological assay; protein-protein interaction;
 KW fluorescence resonance energy transfer system; FRET system.
 XX Synthetic.
 OS Aequorea victoria.
 XX Key Location/Qualifiers
 FT Misc-difference 65
 FT /label= S65A
 FT /note= "Ser to Ala mutation"
 XX WO9806737-A1.
 XX 19-FEB-1998.

PF 15-AUG-1997; 97WO-US014593.
XX
PR 16-AUG-1996; 96US-0024050P.
PR 30-AUG-1996; 96US-00706408.
XX
PA (REGC) UNIV CALIFORNIA.
PA (UYOR-) UNIV OREGON.
XX (AURO-) AURORA BIOSCIENCES.
PI Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
XX WPI; 1998-159454/14.
DR
XX
PT Nucleic acid encoding mutant green fluorescent proteins having longer
PT wavelength emission - used as markers for probes and as components of
PT fluorescent resonant energy transfer systems, also related vectors and
PT transformants.
XX
PS Claim 5; Page; 120pp; English.
XX
CC This sequence is an engineered Aequorea victoria green fluorescent
CC protein (GFP) of the invention, created from the sequence shown in
CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
CC invention, which encode functional, engineered fluorescent proteins (II)
CC having largely the same sequence as Aequorea green fluorescent protein
CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
CC and having fluorescent properties different from GFP. (II) are useful as
CC reporter molecules in immunological or hybridisation assays, for
CC monitoring proteins in cells and detecting induction of transcription.
CC They are also useful as components of FRET (fluorescence resonance energy
CC transfer) systems, e.g. for detecting protein-protein interactions,
CC cleavage of substrates and changes in potential across a membrane, and
CC further for making fluorescent substrates for protein kinase. (II) are
CC easily distinguished from known green and blue fluorescent proteins, so
CC allow simultaneous measurements of at least 2 proteins in a cell. They
CC emit, and are excited, at relatively long wavelengths where phototoxicity
CC and auto-fluorescence are reduced
XX
SQ Sequence 238 AA;
Query Match 99.8%; Score 1267; DB 2; Length 238;
Best Local Similarity 99.6%; Pred. No. 1.3e-122;
Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYGLTLKFKICTTGKLPVPPPTL 60
Db 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYGLTLKFKICTTGKLPVPPPTL 60
Qy 61 VTTFSYGVCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKPEGDTLV 120
Db 61 VTTFSYGVCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKPEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
Qy 181 YYQONTPLDGPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
Db 181 YYQONTPLDGPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
RESULT 4
AAW52333
ID AAW52333 standard; protein; 238 AA.
XX
AC AAW52333;
XX
DT 10-JUL-1998 (first entry)
XX
DE Engineered green fluorescent protein Y66F.
XX
KW Green fluorescent protein; engineered fluorescent protein; mutein;
KW reporter molecule; immunological assay; protein-protein interaction;

fluorescence resonance energy transfer system; FRET system.
XX Synthetic.
OS Aequorea victoria.
XX
FH Key Location/Qualifiers
FT Misc-difference 66 /label= Y66F
FT /note= "Tyr to Phe mutation"
XX
XX WO9806737-A1.
XX
XX 19-FEB-1998.
XX
XX 15-AUG-1997; 97WO-US014593.
XX
XX 16-AUG-1996; 96US-0024050P.
XX 30-AUG-1996; 96US-00706408.
XX
XX (REGC) UNIV CALIFORNIA.
XX (UYOR-) UNIV OREGON.
XX (AURO-) AURORA BIOSCIENCES.
XX
XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
XX WPI; 1998-159454/14.
XX
PT Nucleic acid encoding mutant green fluorescent proteins having longer
PT wavelength emission - used as markers for probes and as components of
PT fluorescent resonant energy transfer systems, also related vectors and
PT transformants.
XX
XX Claim 5; Page; 120pp; English.
XX
CC This sequence is an engineered Aequorea victoria green fluorescent
CC protein (GFP) of the invention, created from the sequence shown in
CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
CC invention, which encode functional, engineered fluorescent proteins (II)
CC having largely the same sequence as Aequorea green fluorescent protein
CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
CC and having fluorescent properties different from GFP. (II) are useful as
CC reporter molecules in immunological or hybridisation assays, for
CC monitoring proteins in cells and detecting induction of transcription.
CC They are also useful as components of FRET (fluorescence resonance energy
CC transfer) systems, e.g. for detecting protein-protein interactions,
CC cleavage of substrates and changes in potential across a membrane, and
CC further for making fluorescent substrates for protein kinase. (II) are
CC easily distinguished from known green and blue fluorescent proteins, so
CC allow simultaneous measurements of at least 2 proteins in a cell. They
CC emit, and are excited, at relatively long wavelengths where phototoxicity
CC and auto-fluorescence are reduced
XX
SQ Sequence 238 AA;
Query Match 99.7%; Score 1266; DB 2; Length 238;
Best Local Similarity 99.6%; Pred. No. 1.7e-122;
Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYGLTLKFKICTTGKLPVPPPTL 60
Db 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYGLTLKFKICTTGKLPVPPPTL 60
Qy 61 VTTFSYGVCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKPEGDTLV 120
Db 61 VTTFSYGVCFSRYPDHMKRHDFFKFSAMPEGYVQORTIFFKDDGNYKTRAEVKPEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
Qy 181 YYQONTPLDGPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
Db 181 YYQONTPLDGPVLLPDNHYLSTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDLYK 238

CC easily distinguished from known green and blue fluorescent proteins, so
 CC allow simultaneous measurements of at least 2 proteins in a cell. They
 CC emit, and are excited, at relatively long wavelengths where phototoxicity
 CC and auto-fluorescence are reduced
 XX
 SQ Sequence 238 AA;

Query Match 99.6%; Score 1265; DB 2; Length 238;
 Best Local Similarity 99.6%; Pred. No. 2.1e-122;
 Matches 237; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYKGLTKLFICTTGKLPVWPPTL 60
 Db 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYKGLTKLFICTTGKLPVWPPTL 60
 Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOORTIFFKDDGNYKTRAEVKEGDTLV 120
 Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOORTIFFKDDGNYKTRAEVKEGDTLV 120
 Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 Qy 181 YYQNTPTLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
 Db 181 YYQNTPTLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 7
 AAW52336
 ID AAW52336 standard; protein; 238 AA.
 XX
 AC AAW52336;
 DT 10-JUL-1998 (first entry)
 XX
 DE Engineered green fluorescent protein.
 XX
 KW Green fluorescent protein; engineered fluorescent protein; mutin;
 KW reporter molecule; immunological assay; protein-protein interaction;
 KW fluorescence resonance energy transfer system; FRET system.
 XX
 OS Synthetic.
 OS Aequorea victoria.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 61
 FT /label= Phe, Tyr, His, Cys
 XX
 PN WO9806737-A1.
 XX
 PD 19-FEB-1998.
 XX
 PF 15-AUG-1997; 97WO-US014593.
 XX
 PR 16-AUG-1996; 96US-0024050P.
 PR 30-AUG-1996; 96US-00706408.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PA (UYOR-) UNIV OREGON.
 XX
 PA (AURO-) AURORA BIOSCIENCES.
 XX
 PI Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
 XX
 DR WPI; 1998-159454/14.
 XX
 PT Nucleic acid encoding mutant green fluorescent proteins having longer
 PT wavelength emission - used as markers for probes and as components of
 PT fluorescent resonant energy transfer systems, also related vectors and
 PT transformants.
 XX
 PS Claim 47; Page; 120pp; English.
 XX

CC This sequence is an engineered Aequorea victoria green fluorescent
 CC protein (GFP) of the invention, created from the sequence shown in
 CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
 CC invention, which encode functional, engineered fluorescent proteins (II)
 CC having largely the same sequence as Aequorea green fluorescent protein
 CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
 CC and having fluorescent properties different from GFP. (II) are useful as
 CC reporter molecules in immunological or hybridisation assays, for
 CC monitoring proteins in cells and detecting induction of transfection.
 CC They are also useful as components of FRET (fluorescence resonance energy
 CC transfer) systems, e.g. for detecting protein-protein interactions,
 CC cleavage of substrates and changes in potential across a membrane, and
 CC further for making fluorescent substrates for protein kinase. (II) are
 CC easily distinguished from known green and blue fluorescent proteins, so
 CC allow simultaneous measurements of at least 2 proteins in a cell. They
 CC emit, and are excited, at relatively long wavelengths where phototoxicity
 CC and auto-fluorescence are reduced
 XX
 SQ Sequence 238 AA;

Query Match 99.6%; Score 1265; DB 2; Length 238;
 Best Local Similarity 99.6%; Pred. No. 2.1e-122;
 Matches 237; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYKGLTKLFICTTGKLPVWPPTL 60
 Db 1 MSKGEELFTAVVILVELDGVNGHKFSVSGEGDVTYKGLTKLFICTTGKLPVWPPTL 60
 Qy 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOORTIFFKDDGNYKTRAEVKEGDTLV 120
 Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOORTIFFKDDGNYKTRAEVKEGDTLV 120
 Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 Qy 181 YYQNTPTLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
 Db 181 YYQNTPTLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

RESULT 8
 AAW52351
 ID AAW52351 standard; protein; 238 AA.
 XX
 AC AAW52351;
 DT 10-JUL-1998 (first entry)
 XX
 DE Engineered green fluorescent protein.
 XX
 KW Green fluorescent protein; engineered fluorescent protein; mutin;
 KW reporter molecule; immunological assay; protein-protein interaction;
 KW fluorescence resonance energy transfer system; FRET system.
 XX
 OS Synthetic.
 OS Aequorea victoria.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 224
 FT /label= His, Asn, Gln, Thr, Phe, Trp, Tyr
 XX
 PN WO9806737-A1.
 XX
 PD 19-FEB-1998.
 XX
 PF 15-AUG-1997; 97WO-US014593.
 XX
 PR 16-AUG-1996; 96US-0024050P.
 PR 30-AUG-1996; 96US-00706408.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PA (UYOR-) UNIV OREGON.

PA (AURO-) AURORA BIOSCIENCES.
 XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
 PI WPI; 1998-159454/14.
 XX
 XX Nucleic acid encoding mutant green fluorescent proteins having longer
 PT wavelength emission - used as markers for probes and as components of
 PT fluorescent resonant energy transfer systems, also related vectors and
 PT transformants.
 XX
 XX Claim 47; Page; 120pp; English.
 XX
 XX This sequence is an engineered Aequorea victoria green fluorescent
 CC protein (GFP) of the invention, created from the sequence shown in
 CC AAW52334. The GFP DNA was mutated to produce the nucleic acids (I) of the
 CC invention, which encode functional, engineered green fluorescent proteins (II)
 CC having largely the same sequence as Aequorea green fluorescent protein
 CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
 CC and having fluorescent properties different from GFP. (II) are useful as
 CC reporter molecules in immunological or hybridisation assays, for
 CC monitoring proteins in cells and detecting induction of transcription.
 CC They are also useful as components of FRET (fluorescence resonance energy
 CC transfer) systems, e.g. for detecting protein-protein interactions,
 CC cleavage of substrates and changes in potential across a membrane, and
 CC further for making fluorescent substrates for protein kinase. (II) are
 CC easily distinguished from known green and blue fluorescent proteins, so
 CC allow simultaneous measurements of at least 2 proteins in a cell. They
 CC emit, and are excited, at relatively long wavelengths where phototoxicity
 CC and auto-fluorescence are reduced
 XX
 XX Sequence 238 AA;
 SQ
 Query Match 99.6%; Score 1265; DB 2; Length 238;
 Best Local Similarity 99.6%; Pred. No. 2.le-122;
 Matches 237; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSKGEELFTAVPILVELDGVNGHKFSVSGEGDVTYGLTLKFTCTTGKLPVWPPTL 60
 Db 1 MSKGEELFTAVPILVELDGVNGHKFSVSGEGDVTYGLTLKFTCTTGKLPVWPPTL 60
 QY 61 VTFPSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNYKTRAEVKEGDTLV 120
 Db 61 VTFPSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNYKTRAEVKEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
 QY 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
 Db 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
 RESULT 9
 AAW52334 standard; protein; 238 AA.
 XX
 XX AAW52334;
 XX
 XX 10-JUL-1998 (first entry)
 DT
 XX Engineered green fluorescent protein Y66W.
 DE
 XX Green fluorescent protein; engineered fluorescent protein; mutein;
 KW reporter molecule; immunological assay; protein-protein interaction;
 KW fluorescence resonance energy transfer system; FRET system.
 XX
 XX Synthetic.
 OS Aequorea victoria.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 66
 FT

FT
 XX
 XX /label= Y66W
 XX /note= "Tyr to Trp mutation"
 FN WO9806737-A1.
 XX
 XX 19-FEB-1998.
 PD
 XX
 XX 15-AUG-1997; 97WO-US014593.
 PF
 XX
 XX 16-AUG-1996; 96US-0024050P.
 PR
 XX 30-AUG-1996; 96US-00706408.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA (UYOR-) UNIV OREGON.
 PA (AURO-) AURORA BIOSCIENCES.
 XX
 XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
 PI WPI; 1998-159454/14.
 DR
 XX
 XX Nucleic acid encoding mutant green fluorescent proteins having longer
 PT wavelength emission - used as markers for probes and as components of
 PT fluorescent resonant energy transfer systems, also related vectors and
 PT transformants.
 XX
 XX Claim 5; Page; 120pp; English.
 PS
 XX
 XX This sequence is an engineered Aequorea victoria green fluorescent
 CC protein (GFP) of the invention, created from the sequence shown in
 CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
 CC invention, which encode functional, engineered green fluorescent proteins (II)
 CC having largely the same sequence as Aequorea green fluorescent protein
 CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
 CC and having fluorescent properties different from GFP. (II) are useful as
 CC reporter molecules in immunological or hybridisation assays, for
 CC monitoring proteins in cells and detecting induction of transcription.
 CC They are also useful as components of FRET (fluorescence resonance energy
 CC transfer) systems, e.g. for detecting protein-protein interactions,
 CC cleavage of substrates and changes in potential across a membrane, and
 CC further for making fluorescent substrates for protein kinase. (II) are
 CC easily distinguished from known green and blue fluorescent proteins, so
 CC allow simultaneous measurements of at least 2 proteins in a cell. They
 CC emit, and are excited, at relatively long wavelengths where phototoxicity
 CC and auto-fluorescence are reduced
 XX
 XX Sequence 238 AA;
 SQ
 Query Match 99.6%; Score 1265; DB 2; Length 238;
 Best Local Similarity 99.6%; Pred. No. 2.le-122;
 Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKGEELFTAVPILVELDGVNGHKFSVSGEGDVTYGLTLKFTCTTGKLPVWPPTL 60
 Db 1 MSKGEELFTAVPILVELDGVNGHKFSVSGEGDVTYGLTLKFTCTTGKLPVWPPTL 60
 QY 61 VTFPSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNYKTRAEVKEGDTLV 120
 Db 61 VTFPSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQORTIFFKDDGNYKTRAEVKEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHNYIIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
 QY 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
 Db 181 YQOQNTPLDGPVLLPDNHYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
 RESULT 10
 AAW52335 standard; protein; 238 AA.
 XX
 XX AAW52335;
 AC

XX 10-JUL-1998 (first entry)
XX Engineered green fluorescent protein.
DE Green fluorescent protein; engineered fluorescent protein; muten;
KW reporter molecule; immunological assay; protein-protein interaction;
KW fluorescence resonance energy transfer system; FRET system.
XX Synthetic.
OS Aequorea victoria.
XX
XX Key Location/Qualifiers
FH Misc-difference 42
FT /label= Cys, Phe, His, Trp, Tyr
FT
XX
XX
XX WO9806737-A1.
XX
XX 19-FEB-1998.
XX
XX 15-AUG-1997; 97WO-US014593.
XX
XX 16-AUG-1996; 96US-0024050P.
XX 30-AUG-1996; 96US-00706408.
XX
XX (REGC) UNIV CALIFORNIA.
PA (UYOR-) UNIV OREGON.
PA (AURO-) AUROREA BIOSCIENCES.
XX
XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
XX WPI; 1998-159454/14.
XX
XX Nucleic acid encoding mutant green fluorescent proteins having longer
PT wavelength emission - used as markers for probes and as components of
PT fluorescent resonant energy transfer systems, also related vectors and
PT transformants.
XX
XX Claim 47; Page; 120pp; English.
XX
XX This sequence is an engineered Aequorea victoria green fluorescent
CC protein (GFP) of the invention, created from the sequence shown in
CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
CC invention, which encode functional, engineered fluorescent proteins (II)
CC having largely the same sequence as Aequorea green fluorescent protein
CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
CC and having fluorescent properties different from GFP. (II) are useful as
CC reporter molecules in immunological or hybridisation assays, for
CC monitoring proteins in cells and detecting induction of transcription.
CC They are also useful as components of FRET (fluorescence resonance energy
CC transfer) systems, e.g. for detecting protein-protein interactions, and
CC cleavage of substrates and changes in potential across a membrane, and
CC further for making fluorescent substrates for protein kinase. (II) are
CC easily distinguished from known green and blue fluorescent proteins, so
CC allow simultaneous measurements of at least 2 proteins in a cell. They
CC emit, and are excited, at relatively long wavelengths where phototoxicity
CC and auto-fluorescence are reduced
XX
XX Sequence 238 AA;
XX
Query Match 99.6%; Score 1265; DB 2; Length 238;
Best Local Similarity 99.6%; Pred. No. 2.1e-122;
Matches 237; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSKGEELFTAVVPIVLVDGNGVGHKFSVSGEGDVTYKLTILKFCITGKLPVPMPTL 60
DB 1 MSKGEELFTAVVPIVLVDGNGVGHKFSVSGEGDVTYKLTILKFCITGKLPVPMPTL 60
QY 61 VTFPSYGVQCFSRYPDMHKRHDFPKSAMPEGYVQOQRTIFFKDDGNYKTRAEVKEGDTLV 120
DB 61 VTFPSYGVQCFSRYPDMHKRHDFPKSAMPEGYVQOQRTIFFKDDGNYKTRAEVKEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNNGHNYIMADKQNGIKVNFKIRHNIEDGVSQVLAD 180

DB 121 NRIELKGIDFKEDGNILGHKLEYNNGHNYIMADKQNGIKVNFKIRHNIEDGVSQVLAD 180
QY 181 YYQONTPILOGPVLLPDNHYLSQTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
DB 181 YYQONTPILOGPVLLPDNHYLSQTQSALSKDPNEKRDMHVLLEFVTAAGITHGMDLYK 238
RESULT 11
AAW52313
ID AAW52331 standard; protein; 238 AA.
XX
AC AAW52331;
XX
DT 10-JUL-1998 (first entry)
XX
DE Engineered green fluorescent protein S65C.
XX
KW Green fluorescent protein; engineered fluorescent protein; muten;
KW reporter molecule; immunological assay; protein-protein interaction;
KW fluorescence resonance energy transfer system; FRET system.
XX
XX Synthetic.
OS Aequorea victoria.
XX
XX Key Location/Qualifiers
FH Misc-difference 65
FT /label= S65C
FT /note= "Ser to Cys mutation"
XX
XX WO9806737-A1.
XX
XX 19-FEB-1998.
XX
XX 15-AUG-1997; 97WO-US014593.
XX
XX 16-AUG-1996; 96US-0024050P.
XX 30-AUG-1996; 96US-00706408.
XX
XX (REGC) UNIV CALIFORNIA.
PA (UYOR-) UNIV OREGON.
PA (AURO-) AUROREA BIOSCIENCES.
XX
XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
XX WPI; 1998-159454/14.
XX
XX Nucleic acid encoding mutant green fluorescent proteins having longer
PT wavelength emission - used as markers for probes and as components of
PT fluorescent resonant energy transfer systems, also related vectors and
PT transformants.
XX
XX Claim 5; Page; 120pp; English.
XX
XX This sequence is an engineered Aequorea victoria green fluorescent
CC protein (GFP) of the invention, created from the sequence shown in
CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
CC invention, which encode functional, engineered fluorescent proteins (II)
CC having largely the same sequence as Aequorea green fluorescent protein
CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
CC and having fluorescent properties different from GFP. (II) are useful as
CC reporter molecules in immunological or hybridisation assays, for
CC monitoring proteins in cells and detecting induction of transcription.
CC They are also useful as components of FRET (fluorescence resonance energy
CC transfer) systems, e.g. for detecting protein-protein interactions, and
CC cleavage of substrates and changes in potential across a membrane, and
CC further for making fluorescent substrates for protein kinase. (II) are
CC easily distinguished from known green and blue fluorescent proteins, so
CC allow simultaneous measurements of at least 2 proteins in a cell. They
CC emit, and are excited, at relatively long wavelengths where phototoxicity
CC and auto-fluorescence are reduced
XX
XX Sequence 238 AA;

Query Match 99.6%; Score 1265; DB 2; Length 238;
 Best Local Similarity 99.6%; Pred. No. 2.1e-122;
 Matches 237; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVVPILVELDGVNGHKFVSVEGEGDVTYKGLTKLTKICTTGKLPVWPPTL 60
 DB 1 MSKGEELFTAVVPILVELDGVNGHKFVSVEGEGDVTYKGLTKLTKICTTGKLPVWPPTL 60

QY 61 VTFYSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQQRITIFFKDDGNYKTRAEVKFEGDTLV 120
 DB 61 VTFYGVQCFSRYPDHMKRHDFFKSAMPEGYVQQRITIFFKDDGNYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFIRHNIEDGSVOLAD 180
 DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFIRHNIEDGSVOLAD 180

QY 181 YVQONTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238
 DB 181 YVQONTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 12
 AAWS2337
 ID AAWS2337 standard; protein; 238 AA.
 XX
 AC AAWS2337;
 XX
 DT 10-JUL-1998 (first entry)
 XX
 DE Engineered green fluorescent protein.
 XX
 KW Green fluorescent protein; engineered fluorescent protein; mutein;
 KW reporter molecule; immunological assay; protein-protein interaction;
 KW fluorescence resonance energy transfer system; FRET system.
 XX
 OS Synthetic.
 OS Aequorea victoria.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 62
 FT /label= Ala, Val, Phe, Ser, Asp, Asn, Gln, Tyr, His, Cys
 XX
 XX W09806737-AL.
 XX
 PD 19-FEB-1998.
 XX
 PF 15-AUG-1997; 97WO-US014593.
 XX
 PR 16-AUG-1996; 96US-0024050P.
 PR 30-AUG-1996; 96US-00706408.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA (UYOR-) UNIV OREGON
 PA (AURO-) AURORA BIOSCIENCES.
 XX
 XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
 XX WPI; 1998-159454/14.

Nucleic acid encoding mutant green fluorescent proteins having longer wavelength emission - used as markers for probes and as components of fluorescent resonant energy transfer systems, also related vectors and transformants.
 XX
 PS Claim 47; Page; 120pp; English.
 XX
 CC This sequence is an engineered Aequorea victoria green fluorescent protein (GFP) of the invention, created from the sequence shown in CC AAWS2313. The GFP DNA was mutated to produce the nucleic acids (I) of CC invention, which encode functional, engineered fluorescent proteins (II) CC having largely the same sequence as Aequorea green fluorescent protein CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)

CC and having fluorescent properties different from GFP. (II) are useful as CC reporter molecules in immunological or hybridisation assays, for CC monitoring proteins in cells and detecting induction of transfection. CC They are also useful as components of FRET (fluorescence resonance energy CC transfer) systems, e.g. for detecting protein-protein interactions, CC cleavage of substrates and changes in potential across a membrane, and CC further for making fluorescent substrates for protein kinase. (II) are CC easily distinguished from known green and blue fluorescent proteins, so CC allow simultaneous measurements of at least 2 proteins in a cell. They CC emit, and are excited, at relatively long wavelengths where phototoxicity CC and auto-fluorescence are reduced
 XX
 SQ Sequence 238 AA;

Query Match 99.6%; Score 1265; DB 2; Length 238;
 Best Local Similarity 99.6%; Pred. No. 2.1e-122;
 Matches 237; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVVPILVELDGVNGHKFVSVEGEGDVTYKGLTKLTKICTTGKLPVWPPTL 60
 DB 1 MSKGEELFTAVVPILVELDGVNGHKFVSVEGEGDVTYKGLTKLTKICTTGKLPVWPPTL 60

QY 61 VTFYSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQQRITIFFKDDGNYKTRAEVKFEGDTLV 120
 DB 61 VTFYGVQCFSRYPDHMKRHDFFKSAMPEGYVQQRITIFFKDDGNYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFIRHNIEDGSVOLAD 180
 DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFIRHNIEDGSVOLAD 180

QY 181 YVQONTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238
 DB 181 YVQONTPIIDGPVLLPDNHYLSTQSALSKDPNEKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 13
 AAWS2349
 ID AAWS2349 standard; protein; 238 AA.
 XX
 AC AAWS2349;
 XX
 DT 10-JUL-1998 (first entry)
 XX
 DE Engineered green fluorescent protein.
 XX
 KW Green fluorescent protein; engineered fluorescent protein; mutein;
 KW reporter molecule; immunological assay; protein-protein interaction;
 KW fluorescence resonance energy transfer system; FRET system.
 XX
 OS Synthetic.
 OS Aequorea victoria.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 220
 FT /label= His, Asn, Gln, Thr
 XX
 XX W09806737-AL.
 XX
 PD 19-FEB-1998.
 XX
 PF 15-AUG-1997; 97WO-US014593.
 XX
 PR 16-AUG-1996; 96US-0024050P.
 PR 30-AUG-1996; 96US-00706408.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA (UYOR-) UNIV OREGON.
 PA (AURO-) AURORA BIOSCIENCES.
 XX
 XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
 XX WPI; 1998-159454/14.

PT Nucleic acid encoding mutant green fluorescent proteins having longer
PT wavelength emission - used as markers for probes and as components of
PT fluorescent resonant energy transfer systems, also related vectors and
PT transformants.

PS Claim 47; Page; 120pp; English.

XX This sequence is an engineered Aequorea victoria green fluorescent
CC protein (GFP) of the invention, created from the sequence shown in
CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
CC invention, which encode functional, engineered fluorescent proteins (II)
CC having largely the same sequence as Aequorea green fluorescent protein
CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
CC and having fluorescent properties different from GFP. (II) are useful as
CC reporter molecules in immunological or hybridisation assays, for
CC monitoring proteins in cells and detecting induction of transcription.
CC They are also useful as components of FRET (fluorescence resonance energy
CC transfer) systems, e.g. for detecting protein-protein interactions,
CC cleavage of substrates and changes in potential across a membrane, and
CC further for making fluorescent substrates for protein kinase. (II) are
CC easily distinguished from known green and blue fluorescent proteins, so
CC allow simultaneous measurements of at least 2 proteins in a cell. They
CC emit, and are excited, at relatively long wavelengths where phototoxicity
CC and auto-fluorescence are reduced

XX Sequence 238 AA;

Query Match 99.6%; Score 1265; DB 2; Length 238;
Best Local Similarity 99.6%; Pred. No. 2.1e-122;
Matches 237; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MSKGEELFTAVVPIVLVDGVDNGHKFSVSGEGDVTYKGLTLKFTCTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVDGVDNGHKFSVSGEGDVTYKGLTLKFTCTTGKLPVWPPTL 60
Qy 61 VTTFSGVQCFSRYPDMHKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDMHKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy 121 NRLEKGIKDFKEDGNILGHKLEYNYNHNYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRLEKGIKDFKEDGNILGHKLEYNYNHNYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YYQNTPTILDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YYQNTPTILDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 14
AAW52346
ID AAW52346 standard; protein; 238 AA.
XX AAW52346;
AC AAW52346;
XX
XX 10-JUL-1998 (first entry)
DT
XX Engineered green fluorescent protein.
DE
XX Green fluorescent protein; engineered fluorescent protein; mutin;
KW reporter molecule; immunological assay; protein-protein interaction;
KW fluorescence resonance energy transfer system; FRET system.
XX
XX Synthetic.
OS Aequorea victoria.

XX Key Location/Qualifiers
FH Misc-difference 167
FT /label= Phe, Tyr, His
XX
XX WO9806737-A1.

XX 19-FEB-1998.

XX

PF 15-AUG-1997; 97WO-US014593.
XX
PR 16-AUG-1996; 96US-0024050P.
PR 30-AUG-1996; 96US-00706408.

XX (REGC) UNIV CALIFORNIA.
PA (UYOR-) UNIV OREGON.
PA (AURO-) AUROA BIOSCIENCES.

XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;
PI WPI; 1998-159454/14.

XX Nucleic acid encoding mutant green fluorescent proteins having longer
XX wavelength emission - used as markers for probes and as components of
PT fluorescent resonant energy transfer systems, also related vectors and
PT transformants.

XX Claim 47; Page; 120pp; English.

XX This sequence is an engineered Aequorea victoria green fluorescent
CC protein (GFP) of the invention, created from the sequence shown in
CC AAW52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
CC invention, which encode functional, engineered fluorescent proteins (II)
CC having largely the same sequence as Aequorea green fluorescent protein
CC (GFP) but differing by at least the substitution T203X (X = H, Y, W or F)
CC and having fluorescent properties different from GFP. (II) are useful as
CC reporter molecules in immunological or hybridisation assays, for
CC monitoring proteins in cells and detecting induction of transcription.
CC They are also useful as components of FRET (fluorescence resonance energy
CC transfer) systems, e.g. for detecting protein-protein interactions,
CC cleavage of substrates and changes in potential across a membrane, and
CC further for making fluorescent substrates for protein kinase. (II) are
CC easily distinguished from known green and blue fluorescent proteins, so
CC allow simultaneous measurements of at least 2 proteins in a cell. They
CC emit, and are excited, at relatively long wavelengths where phototoxicity
CC and auto-fluorescence are reduced

XX Sequence 238 AA;

Query Match 99.6%; Score 1265; DB 2; Length 238;
Best Local Similarity 99.6%; Pred. No. 2.1e-122;
Matches 237; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSKGEELFTAVVPIVLVDGVDNGHKFSVSGEGDVTYKGLTLKFTCTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVVPIVLVDGVDNGHKFSVSGEGDVTYKGLTLKFTCTTGKLPVWPPTL 60
Qy 61 VTTFSGVQCFSRYPDMHKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDMHKRHDFFKSAPEGYVQORTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy 121 NRLEKGIKDFKEDGNILGHKLEYNYNHNYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Db 121 NRLEKGIKDFKEDGNILGHKLEYNYNHNYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 YYQNTPTILDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 YYQNTPTILDGPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 15
AAW52324
ID AAW52324 standard; protein; 238 AA.

XX AAW52324;

XX 10-JUL-1998 (first entry)

XX Engineered green fluorescent protein Y66H.

XX Green fluorescent protein; engineered fluorescent protein; mutin;
KW reporter molecule; immunological assay; protein-protein interaction;

KW fluorescence resonance energy transfer system; FRET system.

OS Synthetic.
OS Aequorea victoria.

XX Key Location/Qualifiers
XX PH /label= Y66H
FT Misc-difference 66
FT /note= "Tyr to His mutation"

XX WO9806737-A1.

XX PN 19-FEB-1998.

XX XX 15-AUG-1997; 97WO-US014593.

XX XX 16-AUG-1996; 96US-0024050P.

XX PR 30-AUG-1996; 96US-00706408.

XX (REGC) UNIV CALIFORNIA.

XX (UYOR-) UNIV OREGON.

XX (AURO-) AURORA BIOSCIENCES.

XX Tsien RY, Cubitt AB, Heim R, Ormo MF, Remington JS;

XX WPI; 1998-159454/14.

XX Nucleic acid encoding mutant green fluorescent proteins having longer

XX wavelength emission - used as markers for probes and as components of

XX fluorescent resonant energy transfer systems, also related vectors and

XX transformants.

XX Claim 5; Page; 120pp; English.

XX This sequence is an engineered Aequorea victoria green fluorescent
XX protein (GFP) of the invention, created from the sequence shown in
XX AA52313. The GFP DNA was mutated to produce the nucleic acids (I) of the
XX invention, which encode functional, engineered fluorescent proteins (II)
XX having largely the same sequence as Aequorea green fluorescent protein
XX (GFP) but differing by at least the substitution T203X (X= H, Y, W or F)
XX and having fluorescent properties different from GFP. (II) are useful as
XX reporter molecules in immunological or hybridisation assays, for
XX monitoring proteins in cells and detecting induction of transcription.
XX They are also useful as components of FRET (fluorescence resonance energy
XX transfer) systems, e.g. for detecting protein-protein interactions,
XX cleavage of substrates and changes in potential across a membrane, and
XX further for making fluorescent substrates for protein kinase. (II) are
XX easily distinguished from known green and blue fluorescent proteins, so
XX allow simultaneous measurements of at least 2 proteins in a cell. They
XX emit, and are excited, at relatively long wavelengths where phototoxicity
XX and auto-fluorescence are reduced

XX Sequence 238 AA;

Query Match 99.6%; Score 1265; DB 2; Length 238;
Best Local Similarity 99.6%; Pred. No. 2.le-122;
Matches 237; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTAVPILVELDGVNGHKFSVSGEGSDVTYGKLTCLKICTTGKLPVWPPTL 60
Db 1 MSKGEELFTAVPILVELDGVNGHKFSVSGEGSDVTYGKLTCLKICTTGKLPVWPPTL 60

QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQORTIFFKDDGNKYKTRAEVKPEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQORTIFFKDDGNKYKTRAEVKPEGDTLV 120

QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNHYIMADKOKNGIKYNFKIRHNIEDGSVOLAD 180
Db 121 NRTELKIDFKEDGNILGHKLEYNNSHNHYIMADKOKNGIKYNFKIRHNIEDGSVOLAD 180

QY 181 YYQNTPTILDGPVLLPNHLYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238
Db 181 YYQNTPTILDGPVLLPNHLYLSTQSALS KDPNEKRDMHVLLEFVTAAGITHGMDELYK 238

Search completed: April 25, 2005, 15:43:35
Job time : 169 secs